

UNIVERSITE DU QUEBEC

THESE

PRESENTEE Á

L'INSTITUT NATIONAL DE LA RECHERCHE SCIENTIFIQUE

COMME EXIGENCE PARTIELLE

DE LA MAITRISE ES SCIENCES (TELECOMMUNICATIONS)

PAR

CHRISTIAN EMOND

B. SC. A. (GENIE ELECTRIQUE)

CODAGE NUMERIQUE A TAUX VARIABLE

DES SIGNAUX DE PAROLE

AVRIL 1981

REMERCIEMENTS

Je désire remercier mon directeur de thèse, le docteur Peter Kabal, professeur visiteur à l'INRS-Télécommunications, qui m'a assisté tout au long du travail de recherche nécessaire à la rédaction du présent ouvrage.

Je remercie également le professeur Paul Mermelstein, chef du groupe de recherche sur la parole à RBN, pour ses conseils judicieux.

Enfin, je me dois de mentionner la compagnie Bell Canada, laquelle a mis sur pied, en collaboration avec l'INRS-Télécom, un programme d'études post-graduées (niveau maîtrise) à l'intention de ses ingénieurs.

C'est en effet ce programme qui a rendu possible pour moi un retour aux études après cinq ans passées à l'emploi de Bell Canada à titre d'ingénieur.

SOMMAIRE

Dans le projet qui suit, nous étudions le comportement du codage à taux variable dans le domaine temporel. L'algorithme de codage procède à partir d'un bloc de N échantillons d'entrée dont elle évalue la fonction de variance et de là, le nombre de bit à être alloué pour chaque échantillon.

Dans la première partie du travail, nous reprenons les sections de la théorie de l'information qui traitent de notre problème, pour ensuite réaliser une simulation "idéale" d'un algorithme de codage à taux variable. Dans la seconde partie, le comportement du codeur est analysé et comparé aux prévisions théoriques. Nous terminons finalement avec un aperçu des tentatives possibles en vue de réaliser une version "pratique" d'un tel codeur.

Nous concluons de ce travail que le seul véritable gain possible dans le domaine du temps est redevable aux stries de la période fondamentale, ce qui rend impossible toute réalisation pratique d'un tel codeur puisque la transmission de l'information nécessaire relativement aux stries demande un taux d'information latérale beaucoup trop élevé. Toute tentative effectuée en vue de réduire ce taux résulte en une annulation complète du gain par rapport au codage à taux fixe.

TABLE DES MATIERES

REMERCIEMENTS.....	ii
SOMMAIRE.....	iii
TABLE DES MATIERES.....	iv
LISTE DES FIGURES ET TABLEAUX.....	vi
 I - INTRODUCTION.....	 1
II - FONDEMENTS THEORIQUES.....	4
2.1 Définition du problème.....	4
2.2 Codage à taux fixe.....	8
2.3 Codage à taux variable.....	9
2.4 Gain du codage à taux variable.....	12
2.5 Vérification pratique de la formule de gain.....	13
2.6 Autre mesure objective: rapport S/B segmentaire.....	19
III- SIMULATION D'UN CODEUR OPTIMAL.....	22
3.1 Introduction.....	22
3.2 Description du système.....	22
3.2.1 Prédiction.....	25
3.2.2 Calcul de variance.....	27
3.2.3 Ajustement du pas de quantification.....	28
3.2.4 Allocation de bits.....	29
3.2.5 Mémoire tampon: capacité et délai.....	32
3.3 Ajustement et optimisation.....	33
3.3.1 Ajustement du pas de quantification.....	34
3.3.2 Allocation de bits.....	39
IV - ANALYSE ET PERFORMANCE.....	47
4.1 Présentation des résultats obtenus.....	47
4.1.1 Comportement en fonction de la mémoire tampon d'entrée.....	48
4.1.2 Variation suivant la constante de temps W.....	49
4.1.3 Variation suivant la plage d'allocation.....	53

4.2	Analyse des résultats.....	55
4.2.1	Commentaires généraux.....	55
4.2.2	Gain du rapport S/B standard.....	56
4.2.3	Gain du rapport S/B segmentaire.....	67
4.3	Jugements subjectifs.....	73
4.3.1	Méthode utilisée.....	73
4.3.2	Résultats des tests.....	74
4.4	Optimisation du rapport S/B segmentaire.....	77
4.5	Dépendance du taux de canal.....	81
V	- ESSAIS DE REALISATIONS PRATIQUES.....	83
5.1	Introduction.....	83
5.2	Codeur "OPTA" modifié.....	84
5.2.1	Limitation sur la capacité des mémoires tampon.....	84
5.2.2	Limitation de l'information latérale.....	86
5.2.3	Temps d'exécution des calculs.....	88
5.2.4	Description globale de l'algorithme OPTA modifié.....	88
5.2.5	Résultats de l'algorithme OPTA modifié.....	90
5.3	Algorithme de Dubnowski et Crochière.....	94
5.3.1	Introduction.....	94
5.3.2	Description de l'algorithme.....	94
5.3.3	Résultats obtenus.....	97
5.3.4	Commentaires sur les résultats.....	98
VI	- CONCLUSIONS.....	104
6.1	Domaine du temps et domaine de fréquences.....	104
6.2	Conclusions globales du projet.....	105
	BIBLIOGRAPHIE.....	110
	APPENDICE A: Gain théorique du codage à taux variable.....	113
	APPENDICE B: Complément sur le rapport S/B segmentaire.....	116
	APPENDICE C: Comparaison de la complexité des méthodes d'allocation.....	122

LISTE DES FIGURES ET TABLEAUX

FIGURES...

2-1	Comportement de K en fonction de b_i (Quantif. optimal gaussien).....	7
2-2	Gain sur le rapport S/B du codage à taux variable en fonction du délai.....	15
2-3	Estimation de la variance du signal de parole utilisant une fenêtre de:	
	a) 160 échantillons (syllabique).....	16
	b) 2 échantillons (instantanée).....	17
3-1	Schéma bloc de la simulation du codeur à taux variable.....	23
3-2	Processus de normalisation du signal d'entrée $X_s(n)$	36
3-3	Densité de probabilité des amplitudes du signal d'entrée normalisé (variance unitaire) pour différentes valeurs de la constante de temps W de la variance.....	37
4-1	Comportement du gain du rapport S/B en fonction du délai.....	50
4-2	Comportement du rapport S/B en fonction de la constante de temps de la variance.....	52
4-3	Evolution de la fonction de variance S_i^2 en fonction du temps (bloc de 256 échantillons).....	59
4-4	Evolution des allocations en fonction du temps.....	60
4-5	Distorsion en fonction du temps (Allocations à taux fixe de 3 bits/éch.).....	61
4-6	Distorsion en fonction du temps (Allocations de la figure 4-5).....	62
5-1	Agencement des mémoires dans le codeur OPTA modifié.....	85
5-2	Distorsion d'un bloc du codeur OPTA modifié correspondant aux figures 4-5 et 4-6 du codeur OPTA idéal (chap IV).....	91

5-3	Comportement du rapport S/B modifié en fonction de la fréquence des mise à jour (m) du pas de quantification et de l'allocation - OPTA modifié.....	92
5-4	Codeur ADPCM à taux variable: Algorithme de Dubnowski et Crochière.....	96
5-5	Allocation en fonction du temps (tampon de 1024 échantillons et $H(b(n-1))=1.01$).....	100
5-6	Variation du pas de quantification d'un paquet à l'autre.....	101
5-7	Variation du taux d'occupation du tampon d'entrée pour chaque paquet.....	102

TABLEAUX...

I	Mesures objectives obtenues suivant différentes densités de probabilité appliquées à l'ajustement du pas de quantification.....	35
II	Résultats comparatifs des méthodes d'allocations.....	42
III	Information sur les quantificateurs uniformes en fonction de la densité utilisée.....	44-45
IV	Vecteur de gain marginal D.....	46
V	Comportement du rapport S/B en fonction de la plage d'allocation.....	54
VI	Vecteur d'allocation empiriquement adapté aux comportements du codeur à taux fixe.....	57
VII	Effets de l'allocation variable sur la déviation standard des rapports S/B.....	71
VIII	Résultats des tests subjectifs.....	75
IX	Comportement en fonction de la capacité de la mémoire tampon d'entrée.....	97
X	Comportement en fonction de la longueur du paquet.....	98
XI	Comportement en fonction du multiplicateur $H(b(n-1))$...	98

I. INTRODUCTION

Le codage numérique de la parole, bien que souhaitable à cause des nombreux avantages qu'il présente, entraîne un débit binaire qui requiert une largeur de bande grandement accrue par rapport à son équivalent analogique.

Depuis de nombreuses années, les chercheurs se sont donc attardés à essayer de réduire le taux de transmission sans affecter la qualité du signal.

Pour ce faire, la majorité des algorithmes de codage s'attachent à réduire la redondance d'information entre les échantillons successifs (ex.:prédiction linéaire) et à combattre la non-stationnarité du signal de parole (ex.:adaptation du pas de quantification).

Outre ces moyens, l'objet du présent ouvrage est d'exploiter une avenue nouvelle - l'allocation variable de bits par échantillon - et d'évaluer le gain qu'elle est en mesure de procurer lorsqu'utilisée en conjonction avec un algorithme de codage utilisant la prédiction linéaire et l'adaptation du pas de quantification.

L'idée de l'allocation variable est séduisante. Elle provient du principe qui veut que lorsque les ressources sont limitées (ex.:coût des canaux de transmission ou des organes de mémoires) il devient souvent préférable de distribuer à chacun suivant ses besoins et non d'une façon "égalitaire".

L'algorithme est donc ici envisagé comme un opérateur agissant sur un "bloc" d'échantillons plus ou moins long pour lequel il dispose d'un nombre limité de bits qu'il doit distribuer à chaque composante suivant un critère prédéterminé.

L'allocation à taux variable ne doit cependant pas être confondue avec le codage utilisant des mots de longueur variable, quelquefois appelé "codage entropique" [15]. Alors que ce dernier utilise au niveau de la quantification une précision uniforme d'un échantillon à l'autre, notre algorithme utilise différents quantificateurs et partant, une précision qui peut varier pour chaque échantillon.

Le codeur que nous avons choisi pour être le "support" de l'algorithme, est de type "ADPCM" ("Adaptive Differential Pulse Code Modulation"), lequel combine un prédicteur fixe et un quantificateur adaptatif. On en trouve une description dans un article général sur le codage de la parole par J.L. Flanagan et al. [22].

Le principe de l'allocation variable est déjà en usage dans les algorithmes de codage par transformée (ATC, codage en sous-bandes) [19,20]. Cette technique n'a cependant que très peu été utilisée dans le domaine du temps. J.J. Dubnowski et R.E. Crochiere [23] ont cependant proposé un algorithme très original pour le codage à taux variable dans le domaine du temps. C'est d'ailleurs cet article qui a attiré notre attention pour finalement nous amener à effectuer le présent ouvrage.

Nous élaborons le principe de l'allocation variable ainsi que le critère qui doit présider aux allocations dans la partie théorique qui forme le chapitre II. Dans le chapitre suivant, nous nous attachons aux simulations ainsi qu'aux modifications apportées à l'algorithme dans le but d'optimiser le gain créé par l'allocation variable de bits par échantillon. Nous étudions ensuite, dans le chapitre IV, le comportement du gain par rapport à certaines variables clé et terminons avec un essai de simulation pratique (chapitre V).

Dans les travaux que nous entreprenons, nous visons à une qualité au moins équivalente à celle d'un lien téléphonique interurbain ("toll quality").

II. FONDEMENTS THEORIQUES

2.1- DEFINITION DU PROBLEME

Avant d'entrer dans les manipulations mathématiques, il serait bon de poser clairement le problème: soit un bloc de N échantillons et un canal de transmission d'une capacité de C bits par échantillon; il s'agit de déterminer la meilleure façon d'attribuer le total de bits disponibles pour le bloc ($N \times C$) étant donné un critère de qualité prédéterminé.

La façon standard de procéder s'appelle le codage à taux fixe et consiste à coder chacun des échantillons avec C bits. L'autre façon, que nous présentons, consiste à attribuer un nombre variable de bits par échantillon jusqu'à épuisement des $N \times C$ bits disponibles pour le bloc considéré. En termes plus rigoureux, si on définit b_i comme le nombre de bits attribué au i ème échantillon, on écrira la contrainte comme étant:

$$2.01 \quad \sum_{i=1}^N b_i = N \cdot C$$

Avant d'entreprendre l'étude théorique proprement dite du codage

à taux variable, il nous faut deux outils fondamentaux. D'abord une mesure objective de l'erreur de quantification, afin de pouvoir évaluer la performance de l'algorithme choisi. Cette mesure est celle qu'on emploie le plus couramment, nous l'appelons l'erreur moyenne carrée. Soit X l'ensemble des signaux à coder, et $\hat{X}$ l'ensemble des signaux après codage, on définit alors d^2 , l'erreur en moyenne carrée, comme:

$$2.02a \quad d^2 = \frac{1}{N} \sum_{i=1}^N (X_i - \hat{X}_i)^2$$

et ensuite une mesure de signal sur bruit (S/B):

$$2.02b \quad S/B = 10 \log \left[\frac{\frac{1}{N} \sum_{i=1}^N X_i^2}{\frac{1}{N} \sum_{i=1}^N d_i^2} \right]$$

L'autre outil dont nous avons besoin apparaît comme évident: en effet, avant d'assigner un nombre plus ou moins grand de bits à un échantillon, il est primordial de connaître, en moyenne tout au moins, quelle en sera l'effet global sur la mesure objective choisie. Une telle question a fait l'objet d'études générales dans ce qu'on appelle la "théorie de l'information" et nous n'élaborerons donc pas. Le sujet est traité en détail par R.G. Gallager (1968) dans le chapitre neuf de son livre [1] et il obtient entr'autre, pour une source gaussienne de variance S^2 (alphabet binaire):

$$2.03a \quad R(d^2) = \frac{1}{2} \log_2 \frac{S^2}{d^2} \quad d^2 < S^2$$

$$2.03b \quad R(d^2) = 0 \quad d^2 \geq S^2$$

$R(d^2)$ représente le "taux" ou encore, le nombre de bits par échantillon nécessaire pour atteindre une distorsion égale ou plus petite que d^2 étant donné un signal de répartition gaussienne à variance S^2 . On note d'ailleurs que si la variance du signal d'entrée est plus petite que l'erreur que l'on est prêt à tolérer, il n'est même plus nécessaire de coder le signal ($R(d^2)=0$).

J. Max en 1960 [4], a mesuré expérimentalement d^2 en fonction du taux $R(d^2)$ à partir de quantificateurs optimaux et il en ressort que l'équation 2.03 ne représente qu'une tendance asymptotique. Ainsi on reformule l'équation 2.03a, cette fois isolée par rapport à d^2 , en y ajoutant un facteur d'ajustement 2^{-2K} , que nous considérerons comme constant, malgré que cela soit moins vrai pour b_i petit (voir figure 2-1):

$$2.04 \quad d_i^2 = S_i^2 \cdot 2^{-2(b_i - K)}$$

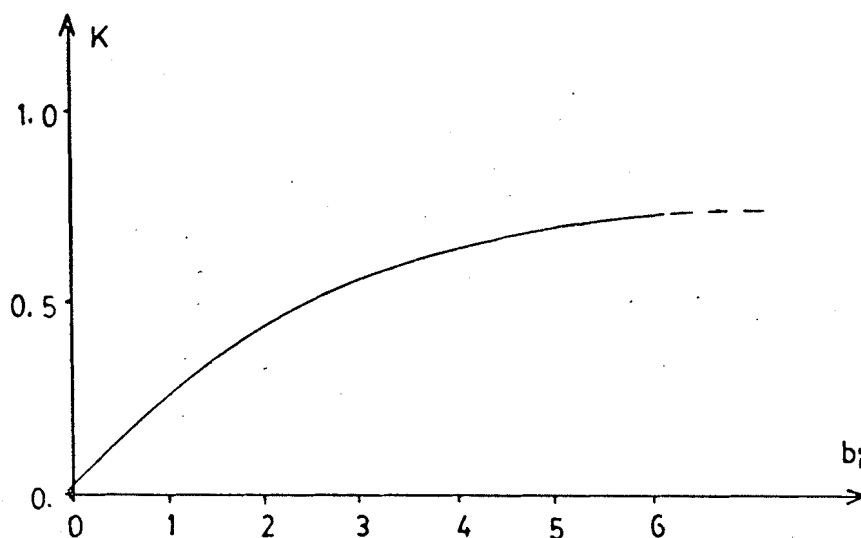


FIGURE 2-1: Comportement de K en fonction de b_i
(Quantif. optimal gaussien)

Notez que les équations 2.02 et 2.04 sont interdépendantes puisqu'une autre mesure sur d^2 amènerait naturellement un autre effet relatif entre d^2 et $R(d^2)$ dans l'équation 2.04.

Avec les équations 2.01, 2.02 et 2.04, nous disposons maintenant des outils suffisants pour entreprendre l'étude comparative des performances du codage à taux fixe et à taux variable. Il est bon de noter toutefois que la valeur de S^2 dans l'équation 2.03 était considérée comme la variance d'un processus stationnaire, alors qu'en pratique on s'intéresse à un estimé local de la variance, indexé $\hat{S}_i^2$. De plus, le processus ayant une moyenne nulle, comme c'est le cas pour la parole, on emploiera indistinctement $\sum_i x_i^2$ ou $\sum_i S_i^2$.

2.2 CODAGE A TAUX FIXE

Dans le cas du codage à taux fixe, l'allocation est fixée à la valeur du taux de canal, soit C bits par échantillon. La distorsion moyenne pour un bloc est donc, suivant l'équation 2.04:

$$2.05 \quad \frac{1}{N} \sum_{i=1}^N d_i^2 = \frac{1}{N} \sum_{i=1}^N S_i^2 2^{-2(C-K)}$$

Le rapport signal/bruit devient ainsi:

$$2.06 \quad S/B = 10 \log \left[\frac{\frac{1}{N} \sum_i S_i^2}{\frac{1}{N} 2^{-2(C-K)} \sum_i S_i^2} \right] \approx 6(C-K)$$

L'équation 2.06 nous ramène au résultat bien connu qui veut qu'en l'absence d'erreur de dépassement, l'addition d'un bit par échantillon représente une amélioration d'environ 6 dB pour le rapport S/B.

2.3 CODAGE A TAUX VARIABLE

Dans le cas du codage à taux variable, le problème est de maximiser le rapport S/B - ou encore minimiser l'erreur (ou distorsion) moyenne carrée - sujet à la contrainte de l'équation 2.01 . Pour trouver ce minimum en combinant la contrainte, on peut utiliser la méthode des multiplicateurs de Lagrange et on forme la fonctionnelle $R(\underline{b}, t)$:

$$2.07 \quad R(\underline{b}, t) = \frac{1}{N} \sum_i S_i^2 \cdot 2^{-2(b_i - K)} + t \cdot \left(\sum_i b_i - N \cdot C \right)$$

On calcule ensuite le différentiel par rapport au multiplicateur t et par rapport à l'allocation b_j dans le but de trouver le minimum:

$$2.08 \quad \frac{dR}{db_j} = \frac{-2}{N} S_j^2 \cdot 2^{-2(b_j - K)} + t = 0.$$

$$2.09 \quad \frac{dR}{dt} = \sum_i b_i - N \cdot C = 0.$$

On résout l'équation 2.08 par rapport à b_j :

$$2.10 \quad b_j = K - \frac{1}{2} \log_2 \left[\frac{N t}{2 S_j^2} \right]$$

On remet la valeur de b_j ainsi trouvée dans l'équation 2.09:

$$2.11 \quad \sum_i b_i = N \cdot C = N \cdot K - \frac{1}{2} \sum_i \log_2 \left[\frac{N t}{2 S_i^2} \right]$$

On transforme la sommation du logarithme en un logarithme du produit:

$$2.12 \quad N \cdot C = N \cdot K - \frac{1}{2} \log_2 \left[\left(\frac{N t}{2} \right)^N \left(\prod_{i=1}^N \frac{1}{S_i^2} \right) \right]$$

On isole par rapport au multiplicateur de Lagrange:

$$2.13 \quad t = \frac{2}{N} \left[\prod_{i=1}^N S_i^2 \right]^{\frac{1}{N}} \cdot 2^{-2(C-K)}$$

Si on remet cette valeur de t dans l'équation 2.10, on obtient finalement:

$$2.14 \quad b_i = C + \frac{1}{2} \log_2 \left(\frac{S_i^2}{d^{*2}} \right) = \frac{1}{2} \log_2 \left[\frac{S_i^2}{2^{2C} d^{*2}} \right]$$

où d^{*2} est donné par la relation:

$$2.15 \quad d^{*2} = \left(\prod_{i=1}^N S_i \right)^{\frac{1}{N}}$$

L'équation 2.14 nous donne donc la formule analytique nous permettant d'effectuer un assignement de bits par échantillon tel que la distorsion carrée est minimisée. Cette distorsion minimum est donnée par:

$$2.16 \quad d_{\min}^2 = d^{*2} 2^{-2(C-K)}$$

Pour trouver l'équation 2.16, on utilise simplement l'équation 2.04 dans laquelle on remplace b_i par sa valeur dans l'équation 2.14.

Ce qui est intéressant de noter, c'est que le niveau de distorsion $d_{\min}^2$ ne dépend plus de l'échantillon et sa variance associée. En d'autres termes, le codeur à taux variable, tel que formulé par les équations ci-haut, attribue les bits de façon à obtenir une distorsion uniforme sur tout le bloc. Par comparaison, le codeur à taux fixe, quant à lui, vise un rapport S/B constant, dont la valeur est donnée par l'équation 2.06.

Dans le codeur à taux variable, la valeur du rapport signal/bruit est donnée par :

$$2.17 \quad S/B = 10 \log \frac{\frac{1}{N} \sum_i S_i^2}{d_{\min}^2} \approx 6(C-K) + 10 \log \left[\frac{\frac{1}{N} \sum_i S_i^2}{(\prod_i S_i^2)^{\frac{1}{N}}} \right]$$

2.4 GAIN DU CODAGE A TAUX VARIABLE

Dans l'équation 2.17, on remarque que le premier terme est semblable au rapport S/B du codeur à taux fixe (section 2.2). L'autre terme correspond donc au gain redevable au codeur à taux variable :

$$2.18 \quad \text{GAIN} = 10 \log \left[\frac{\frac{1}{N} \sum_i S_i^2}{(\prod_i S_i^2)^{\frac{1}{N}}} \right] \geq 0 \quad \forall \{S_i^2\}$$

On remarque en fait que ce gain représente le rapport entre la moyenne arithmétique et la moyenne géométrique de la variance. Si donc le "contour" d'énergie (ou de variance) du signal est relativement stable (ex: un processus stationnaire au sens large), alors le gain est à peu près nul puisque les deux moyennes sont à peu près égales. Par contre si le signal représente un processus non stationnaire (ex: la parole avec des régions de voix et des régions de silences), les deux moyennes diffèrent et un gain s'ensuit. Il

apparaît donc que les non-stationnarités du signal de parole, lesquelles jusqu'ici avaient généré des problèmes au niveau des codeurs, deviennent ici des alliés indispensables au gain !..

On notera en outre que le gain ne dépend pas de la valeur de la constante 2^{-2K} , ni du fait qu'on obtient un accroissement de 6 dB du S/B pour chaque bit additionnelle. Le gain ci-dessus est donc relativement "invariable" à la distribution des signaux d'entrée et on peut voir à l'appendice A qu'un signal de distribution laplacienne obéit à la même formule de gain.

2.5 VERIFICATION PRATIQUE DE LA FORMULE DE GAIN

La première chose qui nous intéresse est de vérifier si le signal de parole contient suffisamment de "non-stationnarités" pour que le gain du codage à taux variable soit potentiellement assez important pour justifier la complexité accrue du codeur.

Nous avons donc conçu un programme sur ordinateur qui calcule une estimation de la variance pour chaque échantillon de voix, suivant une fenêtre exponentielle négative dont la longueur varie suivant la constante α :

$$2.19 \quad S_i^2 = \alpha S_{i-1}^2 + (1-\alpha)X_i^2 \quad W = \frac{1}{(1-\alpha)}$$

A partir des valeurs obtenues pour S_i^2 , on forme les moyennes arithmétiques et géométriques nécessaires au calcul du gain suivant l'équation 2.18. Ces moyennes sont calculées pour un bloc de N échantillons, la valeur de N représentant alors la capacité du tampon d'entrée ou en d'autres termes, le délai minimum qu'il faut prévoir pour le codage.

Le programme a alors été appliqué à cinq fichiers audio contenant des phrases anglaises ou françaises, d'une durée moyenne de 2.4 secondes, et ne contenant pas de long silence. La moyenne des résultats obtenus pour différents délais et différentes valeurs de α apparaît à la figure 2.2.

Les courbes obtenues sont très instructives en ce qui concerne le comportement temporel du signal de parole. Ainsi, pour une variance très syllabique ($W=160$), le signal apparaît comme assez stationnaire jusqu'à environ 100 ms, alors que le gain ne dépasse pas 3 dB. Ceci est prévisible comme résultat puisque la durée moyenne d'un phonème est de 100 ms [7,23]. L'énergie du signal commence donc à varier notablement pour des délais qui dépassent 100 ms (fig 2-3a) permettant ainsi un gain appréciable.

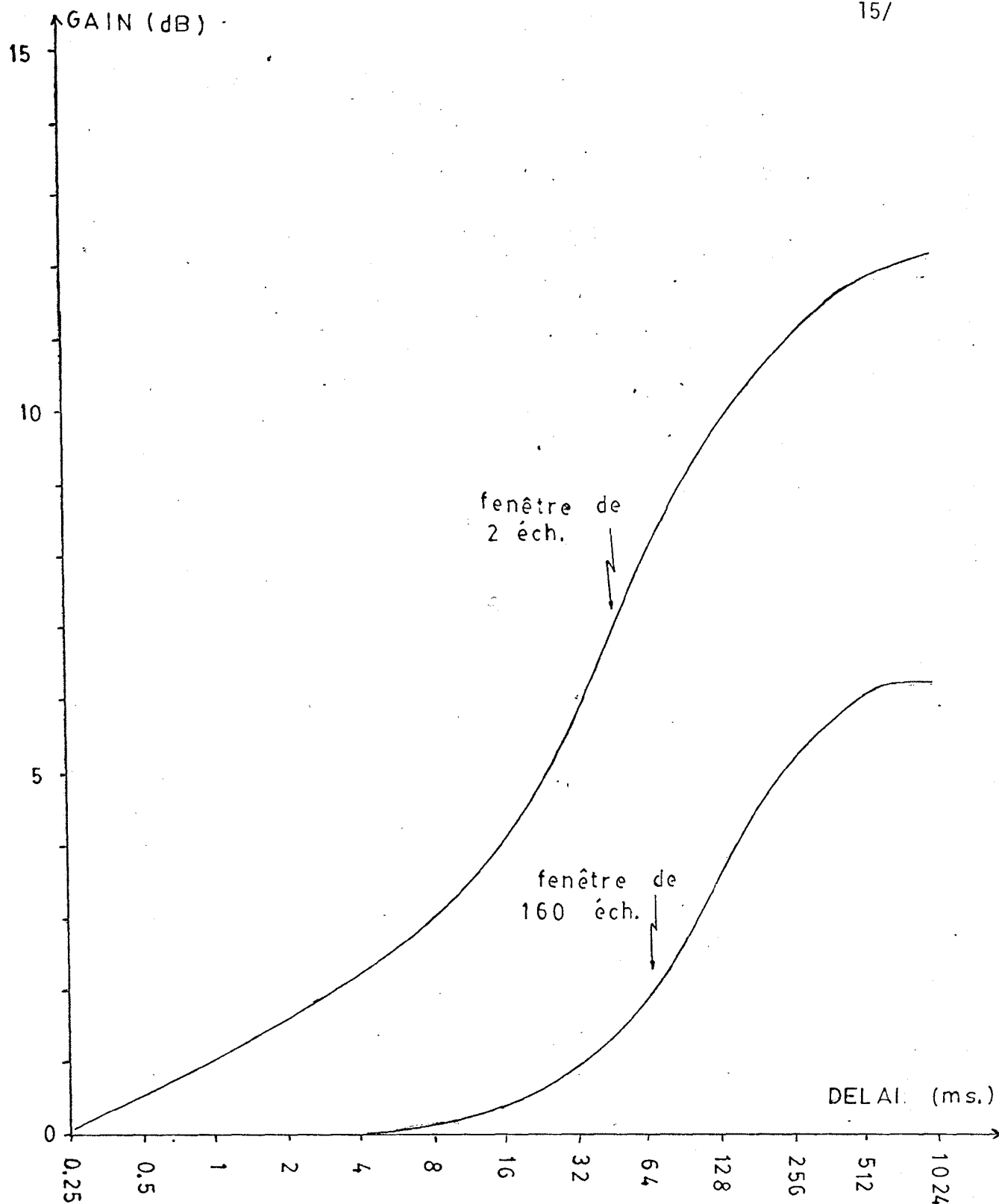


FIGURE 2-2: Gain sur le rapport S/B du codage à taux variable, en fonction du délai (ou longueur du bloc d'échantillons). Fréquence d'échantillonnage de 8000 Hz.

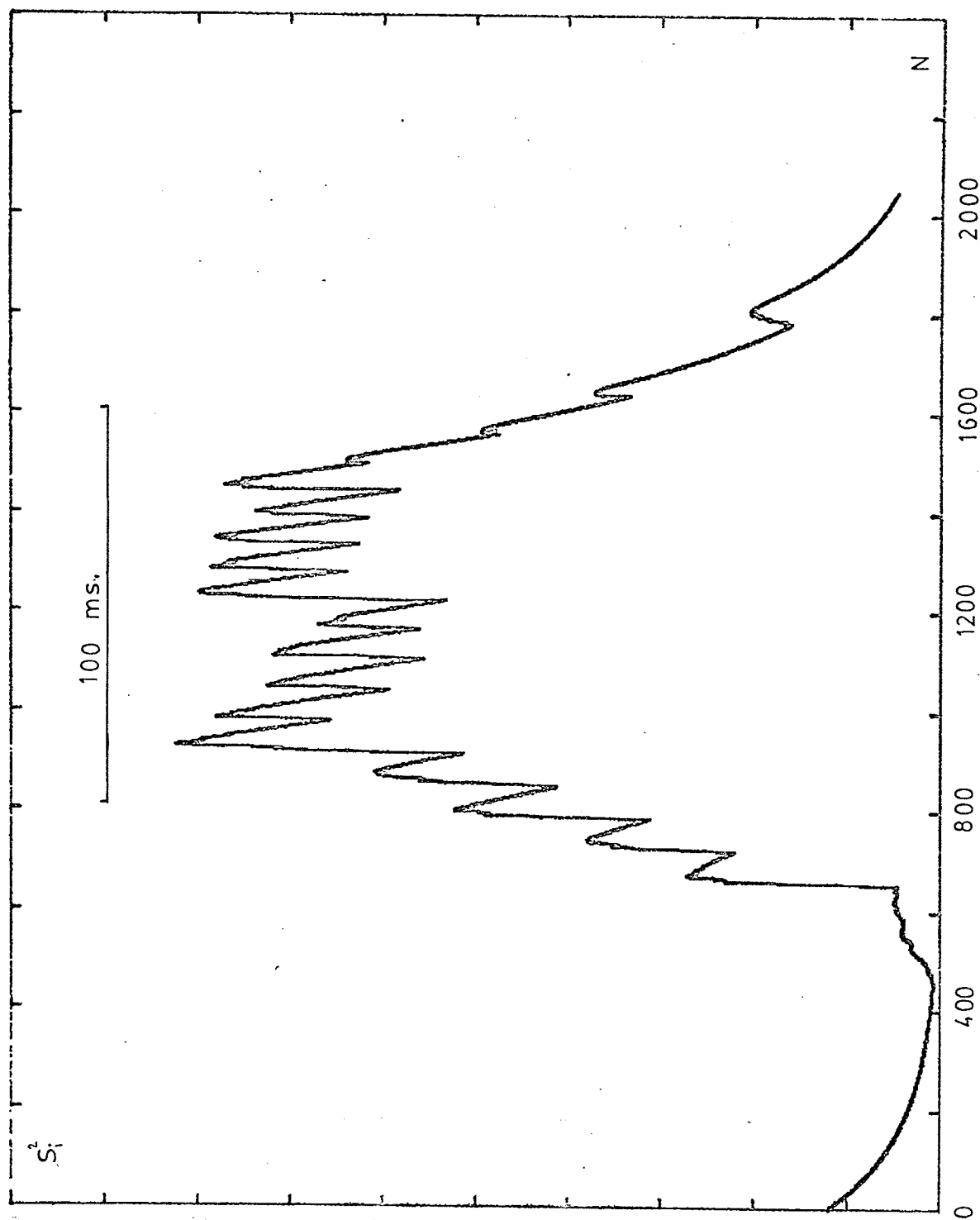


FIGURE 2-3a: Estimation de la variance du signal de parole utilisant une fenêtre de 160 échantillons (variance dite "syllabique").

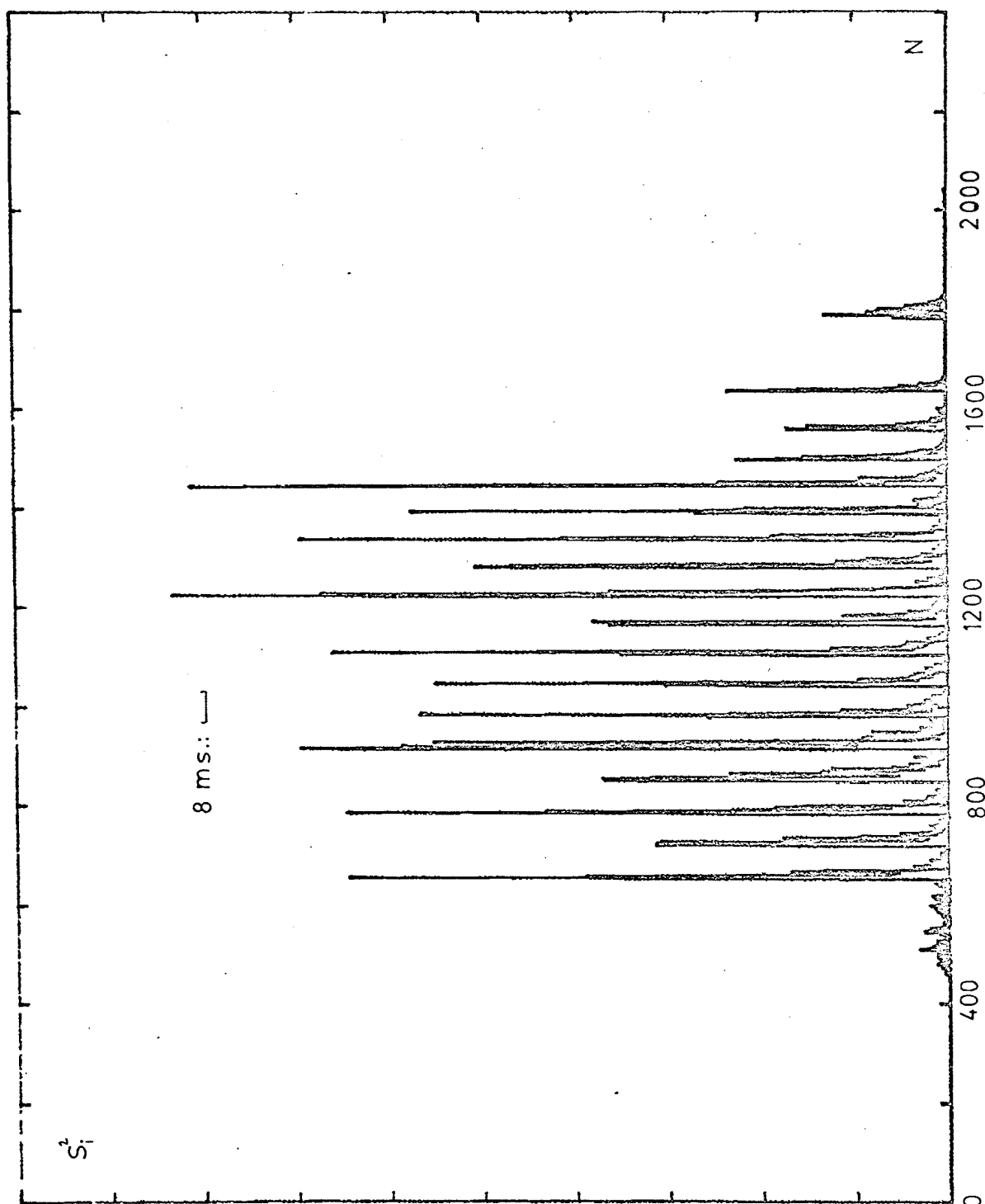


FIGURE 2-3b: Estimation de la variance du signal de parole utilisant une fenêtre de 2 échantillons (variance dite "instantanée")

Lorsque le calcul de variance prend une forme plus "instantanée" ($W < 10$), alors un gain additionnel apparaît, lequel est redevable aux stries de la fondamentale. Ces stries représentent en effet des variations importantes de la variance (fig 2-3b) et comme leur période est courte (moyenne homme-femme: 150 Hz, soit environ 6 ms), ce type de gain se manifeste d'une façon notable dès qu'on atteint un délai de 4 à 8 ms.

On note enfin que le gain n'augmente pas à l'infini comme pourrait le laisser croire la courbe de gain dans l'article de Dubnowski et Crochière (1979) [23]. En effet, un plateau apparaît après un délai de quelque 0.5 seconde. On peut tenter d'expliquer ce phénomène en disant qu'un tel délai contient un ensemble "typique" de phonèmes, de telle sorte qu'en ajoutant d'autres phonèmes, on n'accroît pas la non-stationnarité (ou encore les valeurs relatives des moyennes arithmétiques et géométriques) et ainsi le gain "plafonne". On notera que ces résultats pourraient être modifiés si les phrases contenaient des silences de durée importante.

On peut conclure cette partie en disant que, jusqu'ici, l'idée d'un codeur à taux variable continue d'être intéressante, avec un gain pouvant théoriquement atteindre 12 dB dans le cas de signaux de parole avec variance instantanée.

2.6 AUTRE MESURE SUBJECTIVE: LE RAPPORT S/B SEGMENTAIRE

Dans tous les calculs mathématiques effectués jusqu'ici, c'est le rapport S/B "standard" (c.f. éq. 2.02) qui était optimisé. Dans les expérimentations qui vont suivre, on se sert aussi d'une mesure de S/B dite "segmentaire" à cause du fait qu'on calcule le rapport S/B sur plusieurs segments de voix (typiquement un segment dure 16 msec.) et qu'ensuite, on en fait la moyenne sur toute la phrase. Si on reprend la formulation de M. Nakatsui [33] pour un système fonctionnant à un taux d'échantillonnage de 8000 Hz., on note alors:

$$\begin{aligned}
 2.20 \quad S/B \text{ seg.} &= \frac{1}{M} \sum_{m=0}^{M-1} 10 \log \left[\frac{\sum_{i=1}^{128} S^2(i+128m)}{\sum_{i=1}^{128} d^2(i+128m)} \right] \\
 &= \frac{10}{M} \sum_{m=0}^{M-1} \log(S/B_{128}(m))
 \end{aligned}$$

où M est le nombre de segments de 128 échantillons (16 msec.). La raison pour laquelle cette mesure sera aussi utilisée dans nos résultats est qu'elle a la réputation d'être un meilleur baromètre de l'appréciation subjective.

Les tests effectués par P. Mermelstein [32] et M. Nakatsui [33] rapportent en effet une corrélation beaucoup plus grande entre le jugement subjectif et le S/B segmentaire qu'avec le S/B standard. En

fait, l'article de Nakatsui relève une erreur moyenne (par rapport aux jugements subjectifs) deux fois moindre dans le cas du S/B segmentaire (S/B std.: 7.31 dB, S/B seg.: 4.78 dB).

On peut expliquer la supériorité du S/B segmentaire en disant qu'il est plus voisin du comportement de l'oreille humaine. Cette dernière, en plus d'être sensible à la qualité du signal codé, est aussi affectée par la qualité locale ou les effets de courte durée. A la limite, on pourrait par exemple imaginer une phrase codée d'une façon à peu près parfaite, mais qui comporterait à certains moments une "explosion" de bruit très désagréable. Le S/B standard sera peu affecté par une telle explosion dont les effets seront "absorbés" par l'ensemble de la moyenne (particulièrement si la phrase est longue). Le S/B segmentaire quant à lui, en sera très affecté puisqu'un des segments aura un rapport S/B presque nul. En effet, on peut récrire l'équation 2.20 de la façon suivante:

$$2.21 \quad S/B \text{ seg} = 10 \log \left[\prod_{m=1}^M S/B_m^{(128)} \right]^{1/M}$$

Le S/B segmentaire est donc la moyenne géométrique de tous les rapports S/B individuels et la faiblesse d'un seul élément affecte l'ensemble dans des proportions beaucoup plus notables que pour le S/B standard.

En pratique toutefois, on observe un problème avec ce type de mesure lors de silences. En effet, lorsque l'énergie du signal est à peu près nulle, une légère erreur de quantification résulte en un S/B très négatif, lequel affecte induement l'ensemble. Afin de contourner ce problème, on exclue les segments dont l'énergie totale est moindre de 60 dB par rapport à l'énergie maximum possible [30]. (Dans notre cas, la valeur maximum d'amplitude est 16384 et l'énergie maximum pour un bloc est donc: 128×16384^2 i.e. 3.436×10^{10} . Le seuil est donc fixé à environ 34360).

III.- SIMULATION D'UN CODEUR IDEAL

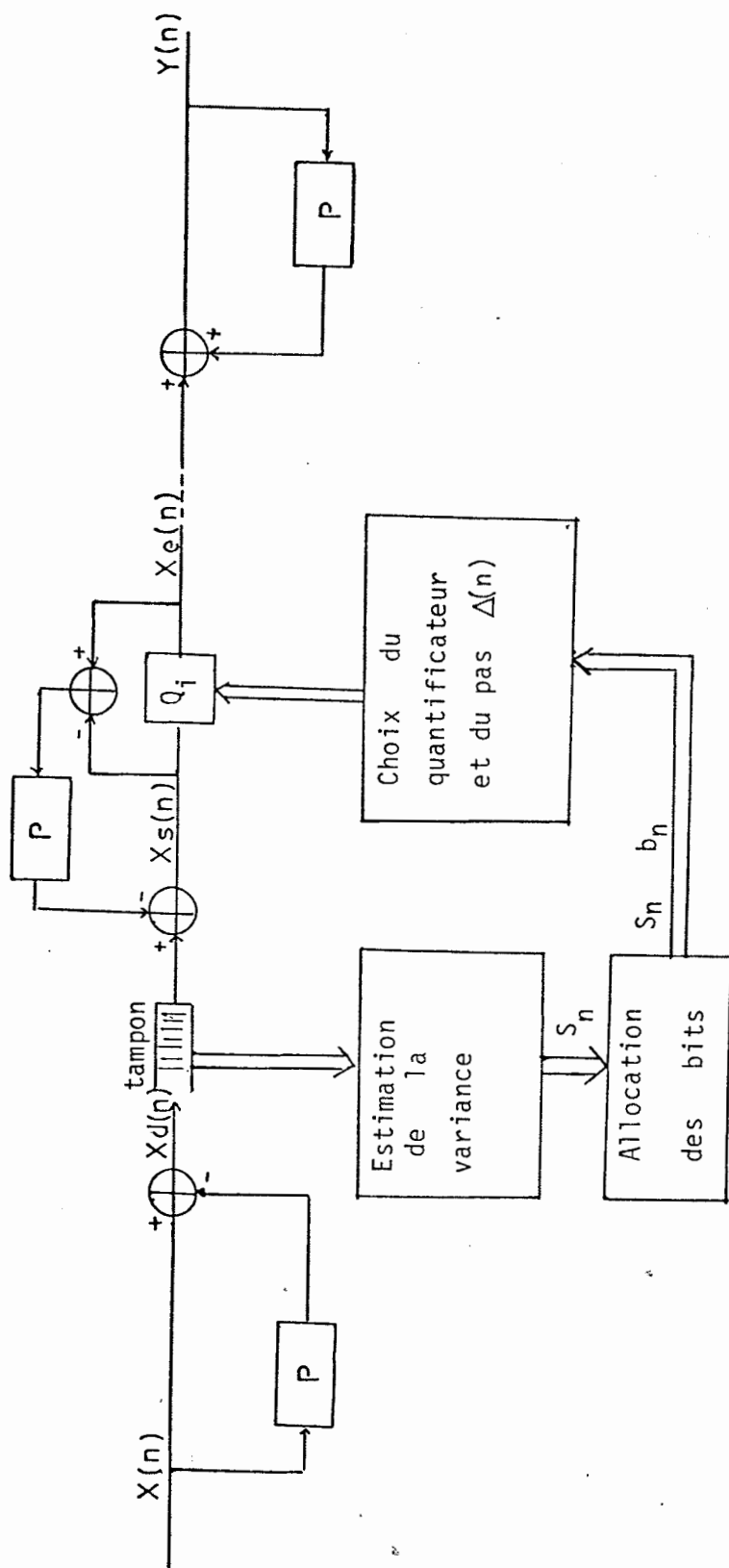
3.1 INTRODUCTION

Ayant couvert dans le chapitre précédent les fondements théoriques du codage à taux variable, nous en sommes maintenant à la vérification des hypothèses par le biais de la simulation sur ordinateur d'un algorithme "idéal".

Le qualificatif "idéal" signifie ici que le codage est effectué sans tenir compte des contraintes d'ordre pratique telles le délai, la complexité des calculs et le taux d'information latérale. Notre but est en effet d'analyser le gain réel du codeur à taux variable dans une situation idéale, pour ensuite considérer la possibilité d'un schéma plus "pratique".

3.2 DESCRIPTION DU SYSTEME

On trouvera à la page suivante (figure 3-1) un schéma d'ensemble du simulateur dont nous élaborerons les diverses composantes dans les



LEGENDE...

Chemin suivi par les
 signaux de parole
 Canaux d'échange d'information

P: Prédicteur

FIGURE 3-1

Schéma bloc de la
 simulation du codeur
 à taux variable.

lignes qui vont suivre.

La simulation est effectuée en utilisant la structure de base d'un codeur "ADPCM" (Adaptive Differential Pulse Code Modulation). On réduit d'abord l'entropie du signal d'entrée $X(n)$ à l'aide d'un prédicteur noté "P". Le signal de différence qui en résulte ($X_d(n)$) est mis en mémoire tampon jusqu'à l'accumulation de N échantillons. Pendant l'arrivée des échantillons, on produit une estimation de variance, de laquelle on dérive l'allocation des bits disponibles. Une fois ces deux vecteurs connus (S_n et b_n) on peut procéder au codage de chaque échantillon en choisissant le quantificateur désiré (en rapport avec l'allocation b_n) et son pas adapté (en rapport avec la variance S_n).

Le signal, une fois quantifié ($X_e(n)$), est reconstruit en utilisant le prédicteur P pour former le signal décodé $Y(n)$. Le codeur n'inclue donc pas certaines considérations telles que l'influence du bruit de canal sur la fiabilité du codeur, le codage binaire, l'influence subjective du délai, les effets d'écho et le contrôle de la mémoire tampon. Il nous apparaît en effet primordial de démontrer d'abord le rendement de l'algorithme de codage, avant de l'envisager -s'il y a lieu- sous des conditions plus réalistes.

Voyons maintenant le détail des composantes du système:

3.2.1 Prédiction

Nous n'entrerons pas ici dans les détails mathématiques entourant le processus de prédiction. Mentionnons toutefois que le prédicteur utilisé est fixe, de deuxième ordre. Les coefficients ont été obtenus de la manière suivante: à l'aide d'un programme ("OPTCOF") nous avons obtenu des coefficients de corrélation $r(1)$ et $r(2)$ pour huit fichiers audio. Nous en avons ensuite fait la moyenne pour obtenir un vecteur de coefficients moyens $r(i)$ (voir lemme 1 dans l'article de Anderson (1975) [14]). On trouve ensuite la valeur des coefficients de prédiction par l'équation vectorielle:

$$3.01 \quad C_i = E(\underline{R})^{-1} \cdot E(\underline{r})$$

où R est la matrice de corrélation, r un vecteur des coefficients de corrélation et C le vecteur des coefficients de prédiction. Les résultats obtenus avec les fichiers audio sont les suivants:

i	$r(i)$	$C(i)$
1	.9189	1.306
2	.7603	-.435

Le gain de prédiction obtenu varie de 7 à 10 dB.

Le stage de prédiction apparaît hors de la boucle de quantification à cause du fait qu'il nous faut une estimation préalable du signal de différence $X_d(n)$ dans le but de calculer la fonction de variance (c.f. fig 3-1) avant même le début du codage proprement dit.

Pour obtenir quand même un signal de sortie $Y(n)$ qui n'intègre pas l'erreur de quantification, on procède avec une technique de boucle de retour du bruit de quantification $q(n)$ telle qu'on l'utilise dans les techniques de modélisation du spectre de l'erreur (Atal et Schroeder 1979 [21], Flanagan et al. 1979 [22]). Suivant le schéma bloc, on a :

$$3.02 \quad Q(z) = X_e(z) - X_s(z)$$

$$3.03 \quad X_s(z) = X_d(z) - P \cdot Q(z)$$

d'où on tire :

$$3.04 \quad X_e(z) = X_d(z) + Q(z) - PQ(z) = X_d(z) + (1-P) \cdot Q(z)$$

De plus on sait que:

$$3.05a \quad Y(z) = \frac{X_e(z)}{(1 - P)}$$

$$3.05b \quad X_d(z) = (1 - P) X(z)$$

En combinant 3.04 et 3.05 on obtient alors:

$$3.06 \quad Y(z) = X(z) + Q(z)$$

qui est le résultat recherché.

3.2.2 Calcul de variance

L'estimation de la fonction de variance s'effectue suivant la même équation que celle mentionnée à la section 2.5, soit:

$$3.07 \quad S_i^2 = \alpha S_{i-1}^2 + (1 - \alpha) X_i^2$$

Mentionnons en plus, que le calcul est effectué d'une façon non causale grâce à l'utilisation de la mémoire tampon qui rend la chose

possible.

3.2.3 Ajustement du pas de quantification

Le codeur dispose d'un nombre de quantificateurs égal au nombre d'allocations possibles du système. Chaque quantificateur est divisé en 2^i pas uniformes centrés à zéro (i est le nombre de bits).

L'ajustement du pas de quantification obéit à la relation de minimisation de la mesure d'erreur, telle que définie par J. Max [4]:

$$3.08 \quad \int_{X_i}^{X_i + \Delta} f(x - y_i) p(x) dx = 0$$

où $f()$ est la relation d'erreur entre le signal d'entrée X et le signal de sortie Y , et $p()$ est la distribution des amplitudes du signal d'entrée. Dans le cas qui nous concerne, on définit $f(x) = x^2$. Quant à $p(x)$, on pourra utiliser différentes distributions et en comparer la performance. Max [4], dans sa publication, fournit un tableau de la valeur du pas, en fonction du nombre de niveaux, pour une distribution gaussienne et une variance unitaire. Pour ce qui concerne les

distributions de type Laplace ou Gamma, on en trouve une liste dans l'article de Paez et Glisson parue en 1972 [18]. Le tableau III à la fin du chapitre donne une liste du pas, de la plage et de l'erreur moyenne carrée en fonction du nombre de bits et de la distribution choisie.

Par exemple, si on utilise une distribution laplacienne avec $S_i^2 = 21500$ et $b_i = 4$, la valeur du pas de quantification serait alors $\Delta_i = 0.456 \times (21500)^{\frac{1}{2}} = 66.9$ et la plage: 501.5.

En pratique, les quantificateurs ont toujours une plage dynamique unitaire (entre -1 et +1). Dans l'algorithme, on divise donc $X_s(n)$ par la valeur de la plage et on quantifie ensuite.

3.2.4 Allocation de bits

Comme nous l'avons vu dans la partie 2.3, l'allocation de bits par échantillon est donnée par la formule analytique de l'équation 2.14.

Toutefois, l'application directe de cette formule conduit à deux problèmes:

- 1) La valeur de b_i peut devenir négative pour $S_i^2 < d^{*2}$

- 2) En pratique, on ne peut allouer qu'un nombre entier de bits. Les erreurs d'arrondissement peuvent avoir comme effet que la somme des b_i n'égale pas exactement le produit $N.C$.

A. Segall (1976) [3] propose deux approches possibles pour ces problèmes. D'abord pour régler le problème (1), il suffit simplement de respecter les formules 2.03a et 2.03b, c'est-à-dire: $b_i = 0$ si $S_i^2 < d^{*2}$. L'équation 2.14 devient alors:

$$3.09a \quad b_i = C + \frac{1}{2} \log_2 \left(\frac{S_i^2}{d^{*2}} \right) \quad S_i^2 > d_{\min}^2$$

$$3.09b \quad b_i = 0 \quad S_i^2 \leq d_{\min}^2$$

Les valeurs de d^{*2} et $d_{\min}^2$ réfèrent ici aux équations 2.15 et 2.16 respectivement.

On notera que l'équation 3.09, en solutionnant le problème (1), accentue le problème (2) puisque, dans certains cas, au lieu de $b_i' = -X$, on aura $b_i = 0$, ce qui aura comme effet de hausser la somme des b_i , $i=[1,N]$.

Dans les algorithmes de codage qui vont suivre, on utilise deux méthodes pratiques pour faire face à ce problème: Dans le premier cas, que nous appellerons la "méthode de la formule analytique" (ou plus simplement: "méthode analytique") nous appliquons simplement la formule 3.09 pour tous les b_i . Ensuite, après en avoir fait la somme, on ajuste la valeur de d^{*2} en fonction du résultat de cette somme et on reprend les allocations avec cette nouvelle valeur de d^{*2} et ce, jusqu'à ce que la somme des allocations satisfasse un critère de précision prédéterminé par rapport aux bits disponibles.

L'autre méthode, celle qui est la plus couramment utilisée dans les algorithmes où l'allocation est variable (ATC, sous-bandes...), est appelée "méthode du gain marginal" (A. Segall [3]), et consiste à calculer le gain marginal, noté:

$$3.10 \quad P_{ij} = S_i^2 (d_j^2 - d_{j-1}^2)$$

où i représente le rang de l'échantillon dans le bloc et j représente la valeur de l'allocation possible. Le terme d_j^2 représente la valeur moyenne de la distorsion, étant donnée une variance unitaire et une allocation de j bits. En d'autres termes, P_{ij} représente le gain qu'on obtient si on alloue j bits plutôt que $j-1$ bits au i ème échantillon. On ordonne ensuite les P_{ij} par ordre décroissant et on alloue ensuite un bit à chaque P_{ij} en partant du plus grand et ce,

jusqu'à épuisement des N.C bits disponibles.

On trouvera dans le tableau IV à la fin du chapitre une liste de vecteur D suivant différents types de distributions (Gauss, Laplace ou Gamma). Ces valeurs sont calculées à partir du tableau III ($E(e^2)$).

L'allocation moyenne, dans les essais qui suivent, est de 3 bits par échantillon (taux du canal), à moins qu'il ne soit spécifié autrement.

On notera en outre que la distribution utilisée pour les amplitudes du signal d'entrée doit être la même pour l'ajustement du pas et l'allocation des bits.

3.2.5 Mémoire tampon - Capacité et délai

Notons ici que la simulation ne tient aucun compte des conditions normales d'utilisation.

On peut voir toutefois sur la figure 3-1 qu'une mémoire tampon apparaît juste après le stage de prédiction. Ceci indique que le

codage ne débutera comme tel qu'après avoir accumulé un total de N échantillons, lesquels sont nécessaires pour établir les allocations individuelles. Le délai augmente donc linéairement avec la longueur du bloc traité.

Il est naturellement nécessaire d'avoir aussi une mémoire tampon à la sortie du transmetteur et à l'entrée du récepteur dans le but de régulariser le flot des bits.

3.3 AJUSTEMENT ET OPTIMISATION

Cette section s'attache à détailler parmi les ajustements faits sur l'algorithme, ceux qui valent d'être mentionnés particulièrement.

On notera qu'aucune tentative de modification n'a été faite en ce qui concerne les coefficients de prédiction et la méthode de calcul de la variance. Cette dernière est en elle-même optimale par sa nature non causale. Nous nous attachons donc plus particulièrement à l'ajustement du pas de quantification et la méthode d'assignement de bits.

3.3.1 Ajustement du pas de quantification

Nous avons comparé ici la performance de l'algorithme à taux fixe utilisant différentes distributions du signal d'entrée pour ajuster le pas du quantificateur. Nous désirons connaître parmi ces distributions laquelle permet le meilleur rapport S/B.

La littérature s'accorde à dire que la distribution du signal d'entrée peut être assimilée à une distribution de type Laplace ou Gamma [7,22]. L'expérience montre que c'est aussi le cas du signal de différence, bien que dans ce cas, la densité Gamma est meilleure à cause de l'accumulation des amplitudes autour de zéro. La plage de quantification est calculée comme étant:

$$3.11 \quad F_i = S_i \cdot A_f(J)$$

où F_i est la plage du i ème échantillon, S_i est la déviation standard et $A_f(J)$ est le vecteur d'ajustement optimal de la plage, étant donné une variance unitaire, et une allocation J (voir tableau III).

Les résultats obtenus à taux fixe (3 bits par échantillon) pour différentes répartitions sont les suivants (moyenne de trois fichiers

audio):

Tableau I - Mesures objectives obtenues suivant différentes densités de probabilité appliquées à l'ajustement du pas de quantification

Variance instantanée...

Répartition	S/B std.	S/B seg.
Gamma	19.5 dB	22.5 dB
Laplace	19.3	22.5
Gauss	21.5	24.7

Variance syllabique...

Répartition	S/B std.	S/B seg.
Gamma	17.6 dB	20.6 dB
Laplace	17.3	20.8
Gauss	15.3	20.6

Au premier abord on peut être surpris des résultats ci-haut, qui démontrent clairement que la répartition gaussienne est la meilleure pour l'algorithme fonctionnant avec une variance instantanée ($N=2$). La raison en est que la quantification est adaptative: le vecteur quantifié n'est donc pas $X_s(i)$ lui-même mais sa valeur normalisée, soit $X_s(i)/F(i) = X_t(i)$ (voir figure 3-2).

En d'autres termes, nous ne cherchons pas la répartition de $X_s(i)$ mais plutôt sa répartition conditionnée à la connaissance de S_i^2 , i.e.: $P(X_s(i)|S_i^2)$. Si on se reporte à la figure 3-2, ceci serait équivalent à évaluer la répartition d'une variable aléatoire notée $X_d(i)/S$. (La répartition de $X_d(i)$ et $X_s(i)$ peut être considérée comme sensiblement la même puisque $X_s(i)$ est en fait le signal $X_d(i)$ auquel on a ajouté le bruit de quantification filtré (c.f. eq. 3.03), ce dernier étant normalement beaucoup plus faible que $X_d(i)$.)

On a donc conçu un programme permettant de relever l'histogramme des amplitudes de $X_d(i)/S_i$ pour trois fichiers audio (deux phrases anglaises et une phrase française). Les résultats obtenus sont montrés à la figure 3-3.

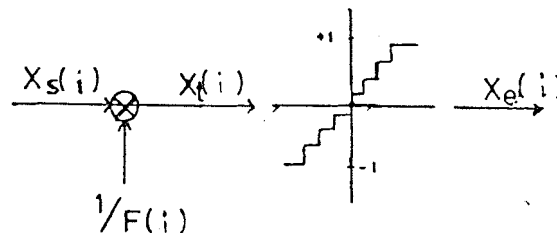


Figure 3-2 Processus de normalisation du signal d'entrée $X_s(i)$

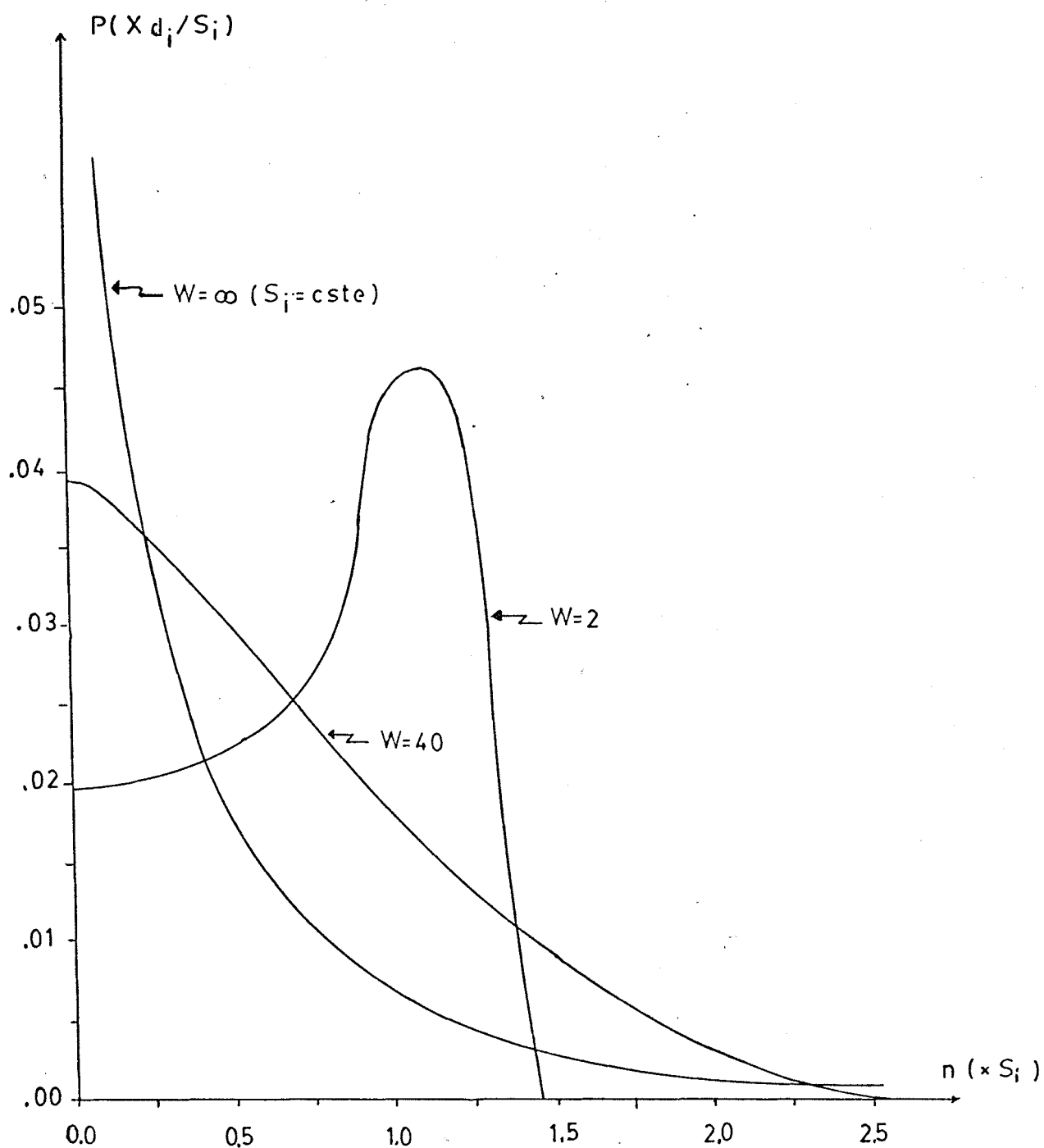


FIGURE 3-3

Densité de probabilité des amplitudes du signal d'entrée normalisé (variance unitaire), pour différentes valeurs de la constante de temps W de la variance.

Comme on peut s'y attendre, la distribution prend une forme bimodale avec une pointe marquée près de $X_d(i) = S_i$ lorsque S_i est calculé avec une constante de temps instantanée. En effet, dans un tel cas, la forme de la fonction de variance ressemble à la forme de la fonction du signal d'entrée, de sorte que la probabilité d'une égalité entre S_i et $X_d(i)$ est très forte. Cette possibilité diminue à mesure que la constante de temps W augmente puisqu'alors, S_i devient une fonction dont les variations se font de plus en plus lentes par rapport à celles de $X_d(i)$, ce qui a pour conséquence que $P(X_d(i)/S_i)$ tend de plus en plus vers $P(X_d(i))$, ce dernier ayant une densité de type Gamma (voir figure 3-3).

Ainsi, on a observé que pour $W < 40$, la densité gaussienne est celle qui s'applique le mieux au comportement de la variable $X_d(i)/S_i$ puisqu'elle implique une moins grande concentration à $X_d(i) = 0$ (si on compare avec Gamma et Laplace), et une dispersion plus faible autour de la déviation standard (ex: $P(X_i = 3S_i)$ vaut 0.0044 dans le cas gaussien, alors que cette probabilité augmente à 0.0102 dans le cas de Laplace, et 0.0113 dans le cas de Gamma).

Dans le cas où la constante d'adaptation de la variance est de type syllabique (nous utiliserons dans les prochaines pages le terme abrégé: "variance syllabique"), la distribution gaussienne est légèrement handicapée au niveau du S/B standard. Toutefois, comme nous le verrons, c'est le rapport S/B segmentaire qui nous intéresse avant tout et ce dernier est équivalent pour les trois types de distributions.

En conclusion, c'est la densité gaussienne que nous retenons pour les prochains travaux et au chapitre IV, tous les résultats qui font l'objet de discussion sont obtenus de l'algorithme utilisant la densité gaussienne.

3.3.2 Allocation de bits

Nous considérons ici l'aspect de l'algorithme utilisé pour l'allocation optimal de bits par échantillon. Nous comparerons donc les deux méthodes sous les aspects suivants:

1) Rapidité d'exécution

Sous l'aspect rapidité, l'algorithme utilisant la formule analytique l'emporte, et de loin. La quantité d'opérations requises est approximativement la suivante (voir appendice C):

3.12a	Formule analytique....	$10 \cdot N \cdot I + 2 \cdot N$
-------	------------------------	----------------------------------

3.12b	Gain marginal.....	$4NC + 3N \log_2 N + \frac{N^2 C}{4}$
-------	--------------------	---------------------------------------

où N est le nombre d'échantillons dans un bloc, I est le nombre moyen d'itérations pour parvenir à une précision raisonnable (5%) et C le taux du canal en bits par échantillon.

A moins d'avoir N petit et I grand (ce qui est typiquement le contraire), on aura avantage à utiliser la formule analytique. Si on prend un cas réaliste avec $C=3$, $I=6$, $N=256$, alors le gain marginal prend environ 42000 opérations contre 15000 pour la formule analytique, soit près de trois fois plus. Ce rapport augmente à 9:1 si $N=1024$ échantillons. Toutefois, au niveau de la simulation, cet handicap a peu d'importance.

2) Capacité de mémoire requise pour les allocations

Sous cet aspect, la méthode utilisant la formule analytique est de nouveau avantageuse puisqu'elle ne nécessite que la mise en mémoire de N valeurs de la fonction de variance et les N valeurs des allocations, pour un total de $2N$.

Dans le cas de la méthode du gain marginal, il faut en plus emmagasiner le vecteur de gain marginal P qui contient N valeurs, pour un total de $3N$, ce qui n'est quand même pas énorme.

3) Possibilité de surcharge de la mémoire tampon de sortie

Au niveau de la simulation, ce problème ne nous préoccupe pas. Mentionnons toutefois que la possibilité de surcharge n'existe que pour la méthode analytique et qu'il faudrait alors une commande de retour pour altérer la valeur de la distorsion d^* en fonction du remplissage du tampon.

4) Efficacité de l'algorithme

Nous avons comparé les résultats en termes de rapport S/B pour les deux méthodes d'allocation:

- la méthode analytique, utilisant l'équation 3.09, considère en fait que la distorsion atteinte correspond au minimum absolu pour un cas idéalement gaussien, étant donné une précision de b , bits [1].

- la méthode de gain marginal est la plus flexible. On peut utiliser ce qu'on veut comme vecteur de gain P_j . Celui qu'on utilisera pour les fins de comparaisons correspond à la distorsion moyenne d'un quantificateur gaussien uniforme [4] (Voir tableau IV sous la rubrique "Gauss").

Les résultats comparatifs correspondent à une moyenne des résultats obtenus sur trois fichiers audio, chacun ayant une plage de

2-6 bits.

Tableau II- Résultats comparatifs des méthodes d'allocation

Variance instantanée...

	S/B std.	S/B seg.
Méthode analytique:	25.9 dB	27.9 dB
Gain marg. gauss. unif.:	25.6	27.4

Variance syllabique...

	S/B std.	S/B seg.
Méthode analytique:	15.7 dB	20.8 dB
Gain marg. gauss. unif.:	14.3	19.5

Les résultats sont à peu près équivalents pour une méthode ou l'autre, probablement parce que la forme de la distribution du signal d'entrée ressemble à une courbe normale.

Dans le cas de la variance syllabique toutefois, il existe un avantage plus net en faveur de la formule analytique. La distribution du signal d'entrée ayant une forme se rapprochant de la distribution gaussienne, il est difficile de prévoir à l'avance quelle sera la méthode la meilleure. Nous avons alors changé le vecteur D pour une forme comparable à la formule analytique, soit :

$$3.13 \quad D: d_j^2 = 2^{-2j}$$

Les résultats obtenus sont de 14.3 et 19.5 dB pour les rapports S/B standard et segmentaire respectivement. Ces résultats sont donc exactement les mêmes qu'avec le vecteur gaussien original (Tableau IV), ce qui montre que l'algorithme ne présente pas une très grande sensibilité à la méthode d'assignement. La méthode analytique conserve toutefois un avantage de quelques dixièmes de décibels vraisemblablement à cause du fait qu'en moyenne elle a une tendance à toujours allouer un peu plus de bits que le nombre disponible. (Cette moyenne est environ de 10 bits pour un total de 768 bits disponibles, ce qui donne un avantage d'environ 0.3 dB si on compte 6 dB/bit). Si on néglige ce détail, on peut dire que les deux méthodes sont équivalentes.

En conclusion de cette sous-section, nous avons sélectionné la méthode de gain marginal puisque cette dernière est plus exacte et permet une plus grande flexibilité.

TABLEAU III

Information sur les quantificateurs uniformes
en fonction de la distribution utilisée.

Allocation	Pas	Plage (F_i)	Erreur $E(e^2)$
Distribution gaussienne			
1	1.596	0.798	0.36340
2	0.996	1.494	0.11880
3	0.586	2.051	0.03744
4	0.335	2.514	0.01154
5	0.1881	2.916	0.00349
6	0.098	3.087	0.00098
Distribution laplacienne			
1	1.414	0.707	0.5000
2	1.087	1.630	0.1963
3	0.731	2.558	0.7170
4	0.456	3.420	0.0254
5	0.281	4.355	0.0087
6	0.170	5.355	0.0028

Tableau III (suite)

Allocation	Pas	Plage (F_i)	Erreur $E(e^2)$
Distribution	Gamma		
1	1.154	0.577	0.6680
2	1.060	1.590	0.3200
3	0.796	2.786	0.1323
4	0.540	4.050	0.0501
5	0.346	5.363	0.0178
6	0.205	6.458	0.0060

Note: Les chiffres correspondant à une allocation de 6 bits ont été obtenus par extrapolation et représentent donc des approximations.

TABLEAU IV

Vecteur de gain marginal D

Type de distribution

Allocation	Gauss	Laplace	Gamma
1	0.6366	0.5000	0.3320
2	0.2446	0.3037	0.3480
3	0.0814	0.1246	0.1877
4	0.0259	0.0463	0.0822
5	0.0081	0.0167	0.0323
6	0.0025	0.0059	0.0118

Note...

Il est intéressant de constater que si l'on utilise le vecteur associé à la densité Gamma, et que l'allocation minimum est de 0 bit, jamais on n'allouera qu'un seul bit à un échantillon puisque $D(2) > D(1)$; (sauf si un échantillon passe de zéro à un bit lors de la toute dernière allocation).

IV - ANALYSE ET PERFORMANCE

Dans ce chapitre, nous sommes maintenant en mesure d'évaluer la performance de l'algorithme de codage à taux variable et d'analyser son comportement.

L'algorithme retenue, tel que décrit dans le chapitre précédent, utilise une distribution gaussienne pour l'ajustement du pas de quantification, ainsi que la méthode du gain marginal pour l'allocation des bits.

4.1 PRESENTATION DES RESULTATS OBTENUS

Nous avons étudié le comportement du codeur en fonction de variables telles la capacité de la mémoire tampon d'entrée, la constante de temps de la variance et finalement la plage d'allocation.

Les courbes représentent une moyenne des résultats obtenus à partir de trois fichiers audio. Le codeur de base présente les caractéristiques suivantes:

- Capacité de mémoire tampon.....256 échantillons
- Constante de temps pour le calcul
de la fonction de variance.....2 échantillons
- Plage d'allocation (taux var.).....2 à 6 bits
- Taux de canal.....3 bits/éch.
- Type d'allocation.....Gain marginal
Vecteur gaussien
min (éq. 3.13)

Dans les pages qui suivent, nous prenons l'une des trois premières caractéristiques que nous faisons varier pour en étudier les effets.

Lorsque le codeur est dit "à taux fixe", cela signifie simplement que l'allocation est fixée à une valeur égale au taux du canal.

4.1.1 Comportement en fonction de la mémoire tampon d'entrée

La capacité de la mémoire tampon d'entrée est une mesure directe du délai qui serait requis lors d'une implantation pratique de

l'algorithme. La figure 4-1 nous montre une comparaison de gains théoriques, S/B standard et segmentaire.

La courbe théorique a été obtenue de la même façon que celle de la figure 2-2, avec cette fois les trois fichiers audio qui ont été utilisés pour les autres courbes de ce chapitre.

La courbe du rapport S/B standard, bien qu'atteignant un gain sensiblement moindre que la courbe théorique, est quand même monotone croissante en fonction de la capacité de mémoire. On observe qu'au départ (entre 0 et 4 msec.) la courbe de gain S/B standard suit de près la courbe de gain théorique. On observe en plus un "plateau" dans la région qui va de 64 à 512 échantillons (8 à 64 msec.).

Quant au rapport S/B segmentaire, on observe une augmentation nette du gain au début, laquelle "plafonne" lorsqu'on augmente la capacité au-delà de 64 échantillons (8 msec.). On constate surtout que le gain maximum est très faible si on le compare au gain théorique.

4.1.2 Variation suivant la constante de temps N

Nous avons mesuré la performance de l'algorithme en fonction de la

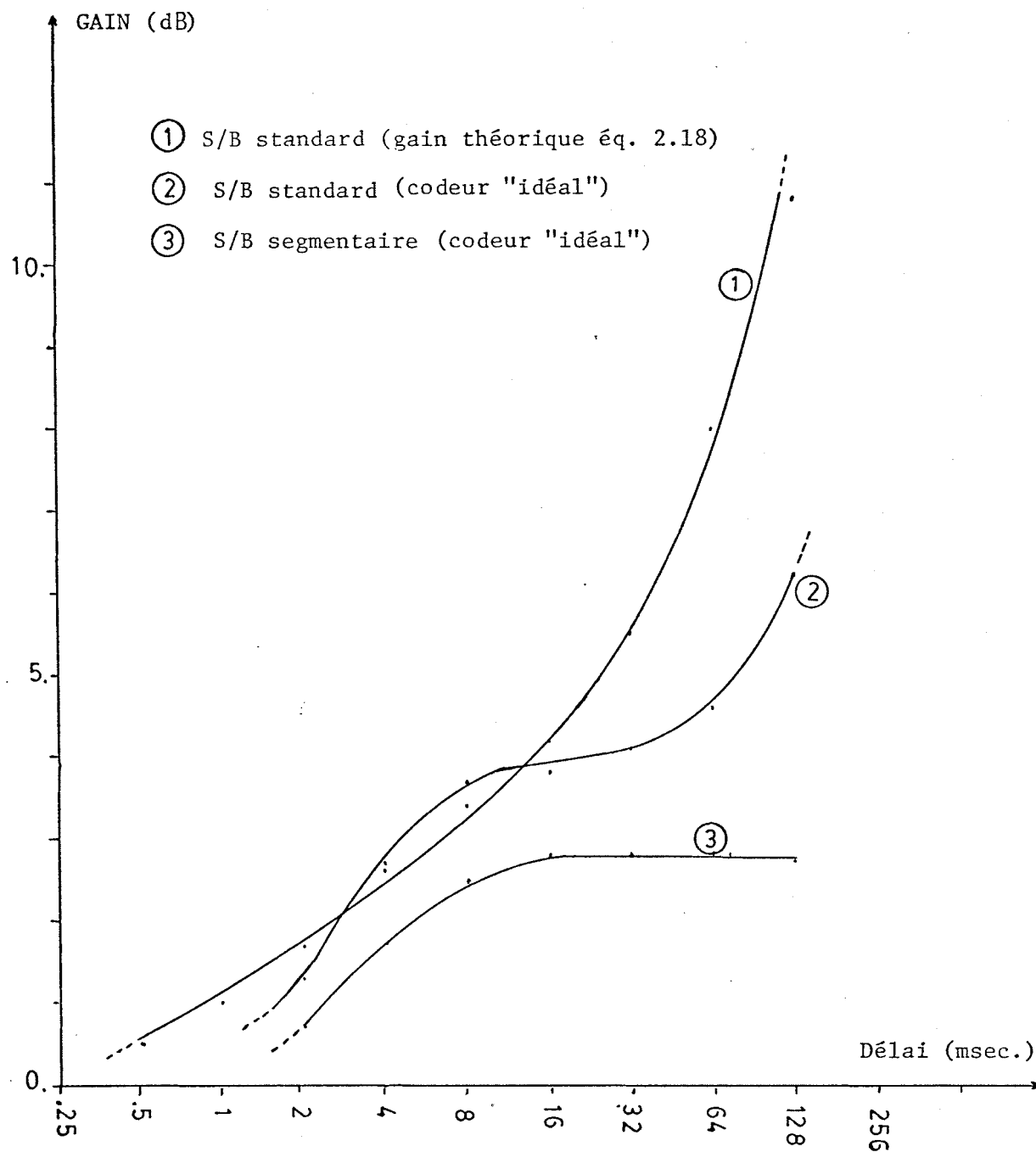


Fig 4-1 Comportement du gain du rapport S/B en fonction du délai.

constante de temps W servant au calcul de la variance. Le rapport entre W et α est décrit à l'équation 2.19.

La figure 4-2 montre la variation des rapports S/B en fonction de la constante de temps W . Par souci de référence, nous avons inclus les rapports S/B d'un codeur ADPCM-F standard utilisant les multiplicateurs de Jayant [8,10]. Un tel codeur utilise le niveau de quantification de l'échantillon précédent pour déterminer la valeur du pas de quantification présent selon la relation:

$$4.01 \quad \Delta_i = \Delta_{i-1} \cdot M(|L_{i-1}|)$$

où Δ_i est le pas de quantification du i ème échantillon, $M()$ est le vecteur des multiplicateurs et enfin, L est le niveau utilisé pour la quantification du i ème échantillon.

Comme on peut le constater sur la figure, seule la variance de type instantanée permet un gain au-delà du même codeur fonctionnant à taux fixe.

Ainsi, pour $W < 10$, l'algorithme à taux variable est supérieure. Par contre, pour $W > 10$, non seulement le gain disparaît, mais il devient

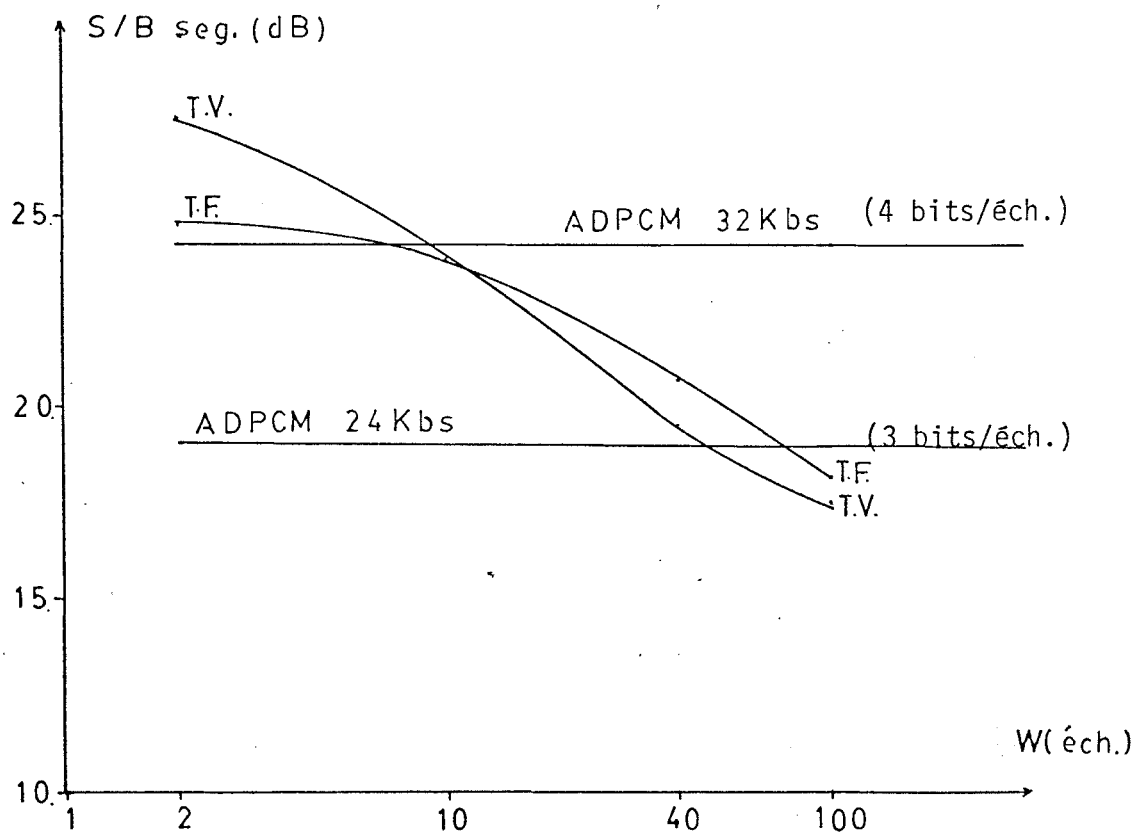
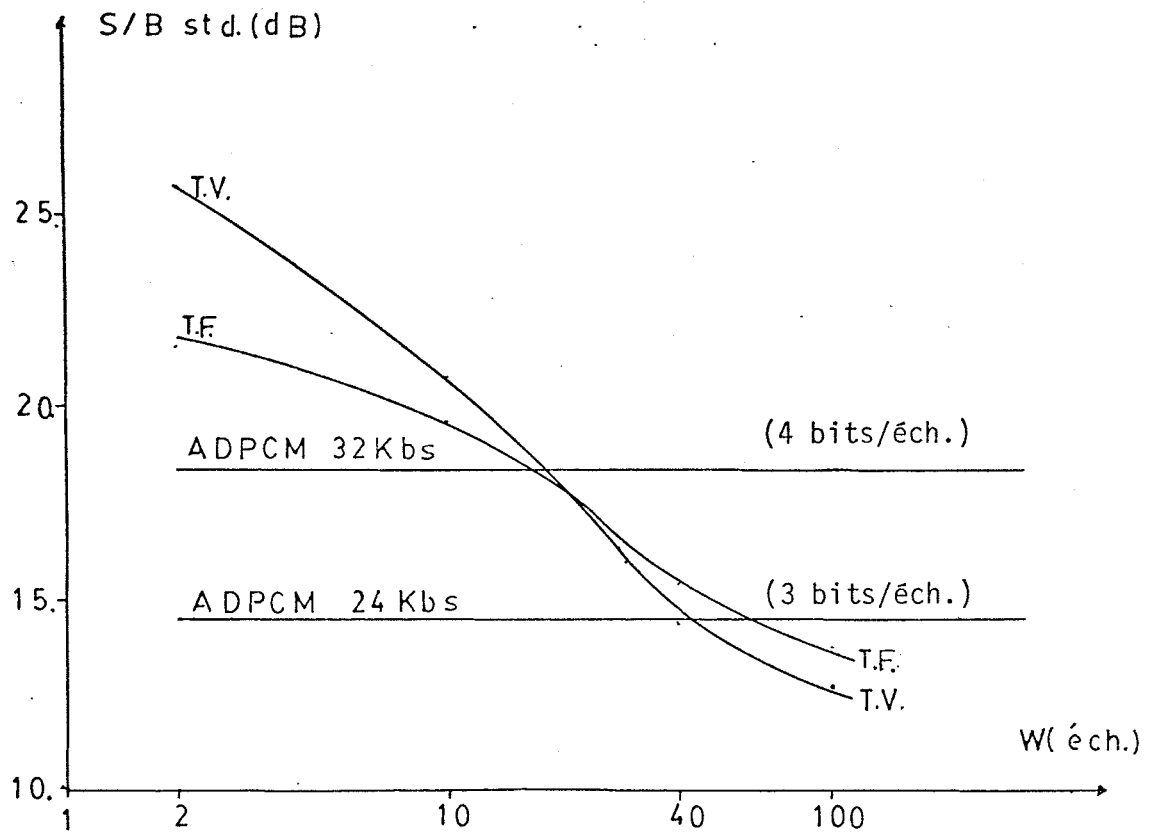


FIG 4-2 Comportement du rapport S/B en fonction de la constante de temps de la variance

moindre par rapport au même codeur fonctionnant à taux fixe.

Le codeur à taux fixe que nous utilisons est supérieur à celui qui utilise les multiplicateurs de Jayant. Ce gain est redevable à la technique utilisée pour l'ajustement du pas de quantification: grâce au tampon d'entrée, il nous est possible de faire une estimation non-causale de la variance en vue d'ajuster le pas de quantification ("forward estimate"), alors que le codeur ADPCM-F standard utilise une estimation en aval ("backward estimate"). Le prix que nous devons "payer" pour ce gain additionnel est le délai ainsi que l'information latérale qui doit alors être transmise séparément.

4.1.3 Variation suivant la plage d'allocation

La figure 4-3 présente le comportement du codeur en fonction de la plage d'allocation, en partant d'une plage nulle (i.e. taux fixe de 3 bits par échantillon) jusqu'à une plage maximum de 7 bits (i.e. de 0 bit à 6 bits par échantillon).

Nous nous en sommes tenu à un ensemble de six plages possibles et avons éliminé les autres par souci de limiter les travaux à un ensemble typique et représentatif.

On notera que le gain maximum est atteint à toutes fins pratiques avec une plage variant entre 2 et 4 bits par échantillon, et qu'elle a même une tendance à diminuer si on diminue la borne d'assignement inférieure.

Tableau V - Comportement du rapport S/B en fonction
de la plage d'allocation

Plage	S/B std.	S/B seg.
3 - 3	21.5 dB	24.7 dB
2 - 4	24.5	26.8
2 - 5	25.6	27.4
2 - 6	25.6	27.5
1 - 6	25.3	27.1
0 - 6	24.9	27.0

On peut expliquer ce phénomène par le fait que le gain du rapport S/B (en dB) n'est pas linéaire comme le prétend le vecteur de gain marginal. En effet, si on se situe dans la région des assignements faibles (un ou deux bits), le rapport S/B du codeur augmente d'environ neuf décibels pour chaque bit ajouté, alors que dans la région où l'allocation est haute (cinq ou six bits), le gain marginal est d'environ cinq décibels. En somme, il semble préférable d'allouer à chaque échantillon un "montant" minimum de départ (dans notre cas:

deux bits par échantillon) et allouer le reste d'une façon variable. En faisant cela, on diminue les allocations de six bits, lesquelles n'apportent souvent qu'un gain marginal assez faible, et représentent des rapports S/B dont l'oreille humaine n'est pas en mesure d'apprécier de toute manière (en effet, lorsque le rapport S/B dépasse 30 dB, l'oreille humaine n'est plus en mesure d'apprécier les variations de qualité [32]).

4.2 ANALYSE DES RESULTATS

4.2.1 Commentaires généraux

On peut dire dans l'ensemble que les résultats de simulations pratiques déçoivent et étonnent à la fois. On aurait pu s'attendre à mieux, au moins en ce qui concerne le rapport S/B standard. Dans cette section, nous allons tenter d'expliquer les phénomènes qui concourent à diminuer sensiblement le gain par rapport à ce qu'il aurait pu être.

Mentionnons toutefois que jusqu'ici, le gain du codage à taux variable n'est possible qu'avec une variance de type instantané. De plus, si on s'attarde au rapport S/B segmentaire, un délai de 256 échantillons (32 msec.) est amplement suffisant pour obtenir tout le

gain possible, ce qui va à l'encontre de ce qu'on prévoyait précédemment [23].

4.2.2 Gain du rapport S/B standard

Alors qu'on aurait pu s'attendre à un gain de près de 12 dB en théorie, on dépasse difficilement 5 dB de gain en pratique.

Lorsque ces résultats nous sont apparus pour la première fois, nous en avons cherché la raison sur plusieurs plans possibles:

Hypothèse #1: Méthode d'allocation non appropriée aux statistiques du signal.

Nous avons essayé plusieurs autres vecteurs de gain marginal en plus des deux dont nous faisons mention au chapitre précédent. A chaque fois, nous obtenions des résultats équivalents ou moindres que ceux que nous utilisons.

A ce sujet, mentionnons le suivant: nous avons fait fonctionner le codeur à taux fixe pour un taux de canal variant successivement de un à six bits par échantillon. Le S/B ainsi obtenu (en dB) a été

transformé en rapport absolu (ex.: 6. dB S/B équivalant à un rapport de 0.25). En faisant cela pour chacun des taux, on a obtenu un nouveau vecteur D adapté au comportement réel du codeur. Le tableau suivant montre les résultats:

Tableau VI - Vecteur d'allocation empiriquement adapté au comportement du codeur à taux fixe.

Bit	S/B std.(dB)	S/B std.(abs.)	D
1	8.9	.12900	.87100
2	17.4	.01820	.11100
3	21.5	.00710	.01100
4	26.1	.00250	.00460
5	31.3	.00074	.00176
6	36.8	.00021	.00053

Les résultats obtenus avec ce vecteur "adapté" sont de 25.4 dB et 27.2 dB pour les rapport S/B standard et segmentaire respectivement. Ces résultats sont donc comparables à ce qu'on obtenait avec un vecteur gaussien. Ce n'est donc pas du côté de la méthode d'allocation qu'il faut chercher. D'ailleurs, il nous est apparu que la méthode est peu sensible au type de vecteur de gain marginal utilisé.

Hypothèse #2: Irréalisme de la formule de gain théorique.

Dans le codage à taux fixe "idéal" le rapport S/B est constant d'un échantillon à l'autre. Chacun sait que cet "idéal" n'est pas réalisable en pratique et qu'on ne peut qu'espérer pouvoir s'en approcher un peu.

Dans le codage à taux variable "idéal", la distorsion d^2 est constante d'un échantillon à l'autre (c.f. éq. 2.15) et correspond à la moyenne géométrique des éléments $\{S_i^2\}$ de la fonction de variance.

Pour vérifier jusqu'à quel point la distorsion devient constante, nous avons étudié en détail le comportement du codeur sur un bloc particulier. Référons-nous pour cela aux figures 4-3 à 4-6 dans les pages qui suivent. On peut voir sur la figure 4-3 d'abord la variance en fonction du temps (bloc de 256 échantillons). L'énergie est surtout localisée entre le 100ième et le 256ième échantillon.

A la figure 4-4, on voit les allocations qui y correspondent. Dans les régions de faible énergie, les allocations oscillent entre 2 et 3 bits, alors que dans la région à haute énergie, les allocations sont supérieures au taux de canal (5 ou 6 bits).

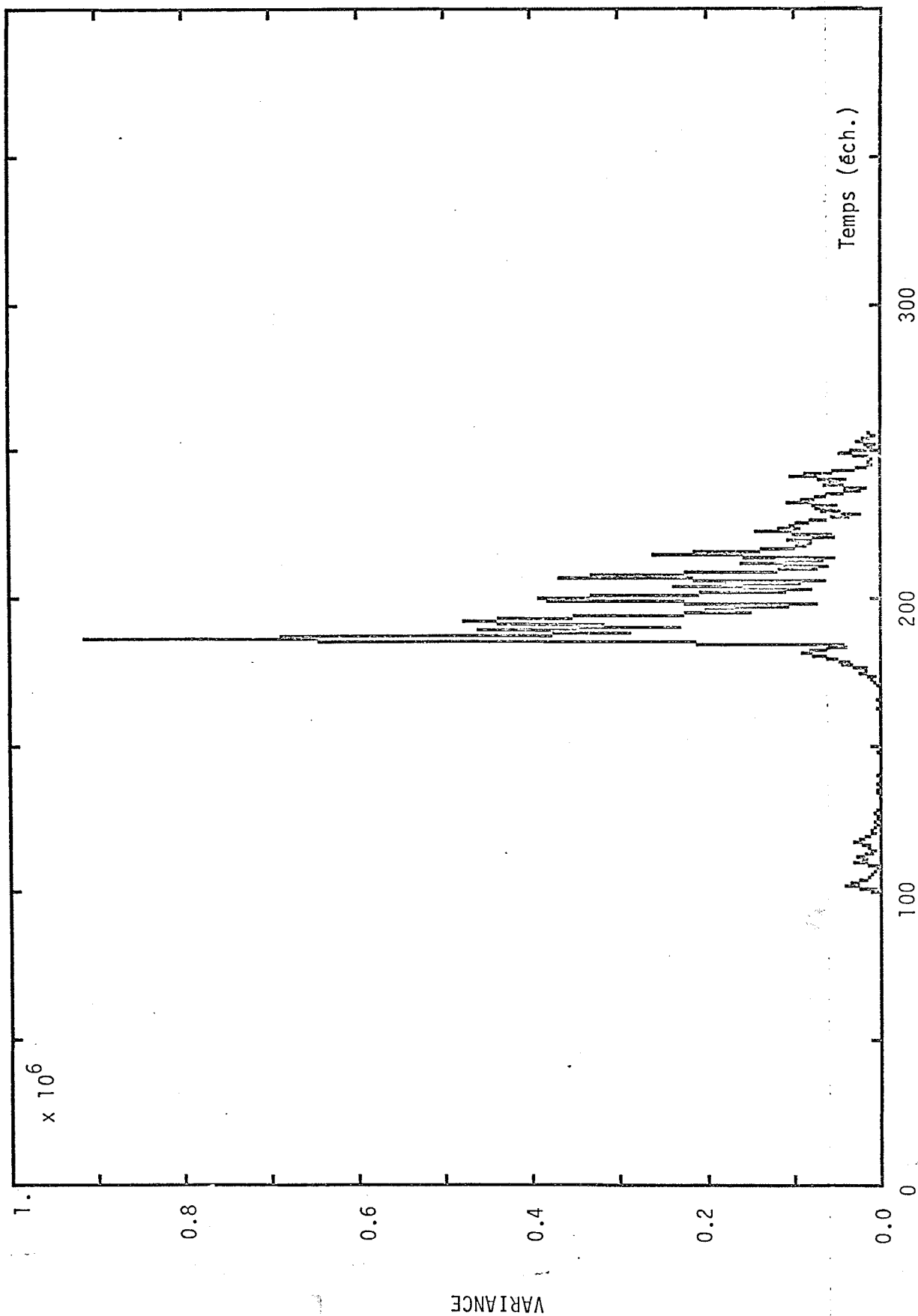


FIG 4-3 Evolution de la fonction de variance S_i^2 en fonction du temps
(Bloc de 256 échantillons) -

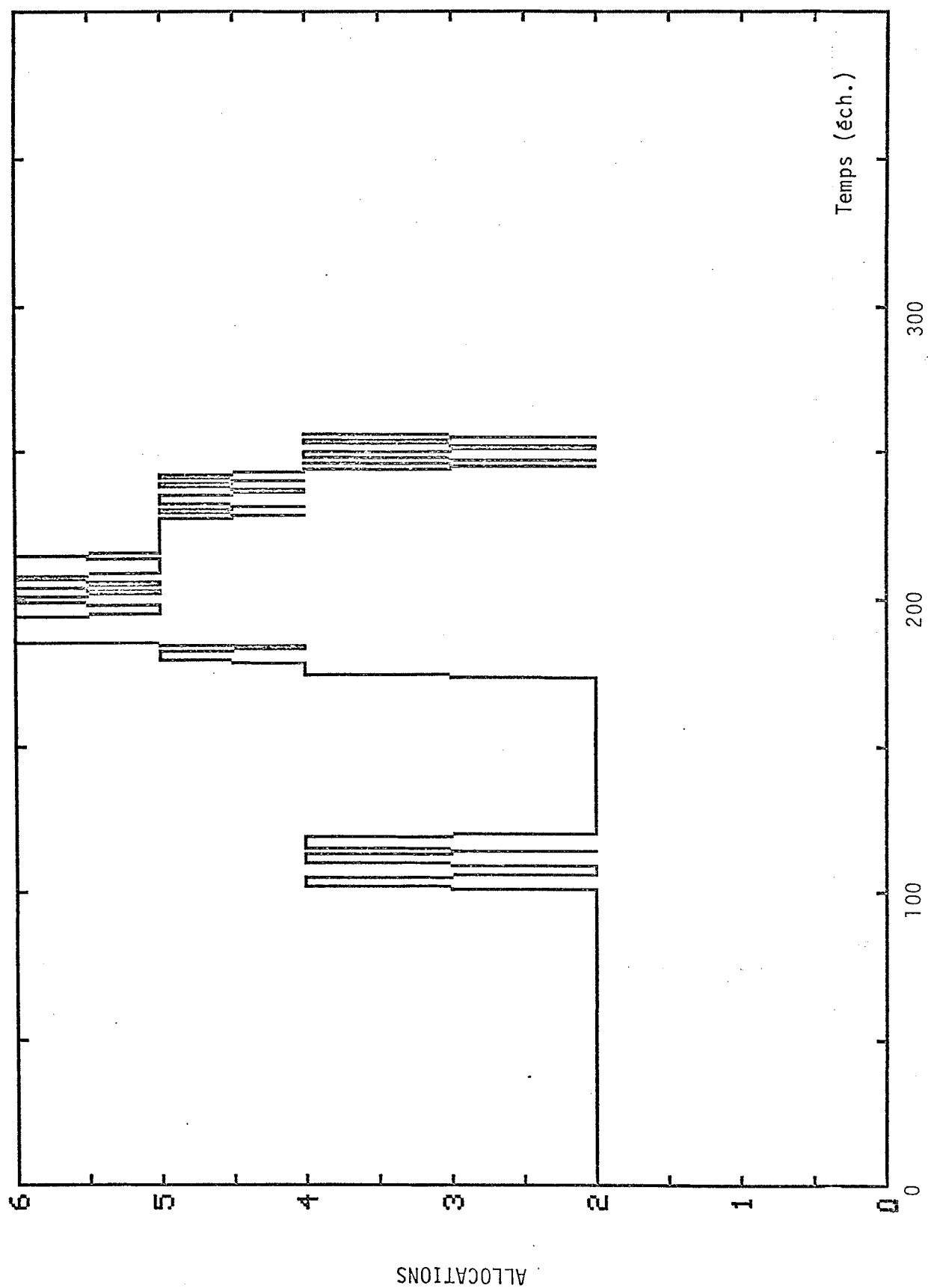


FIG 4-4 Evolution des allocations en fonction du temps (Variance de la fig 4-3).

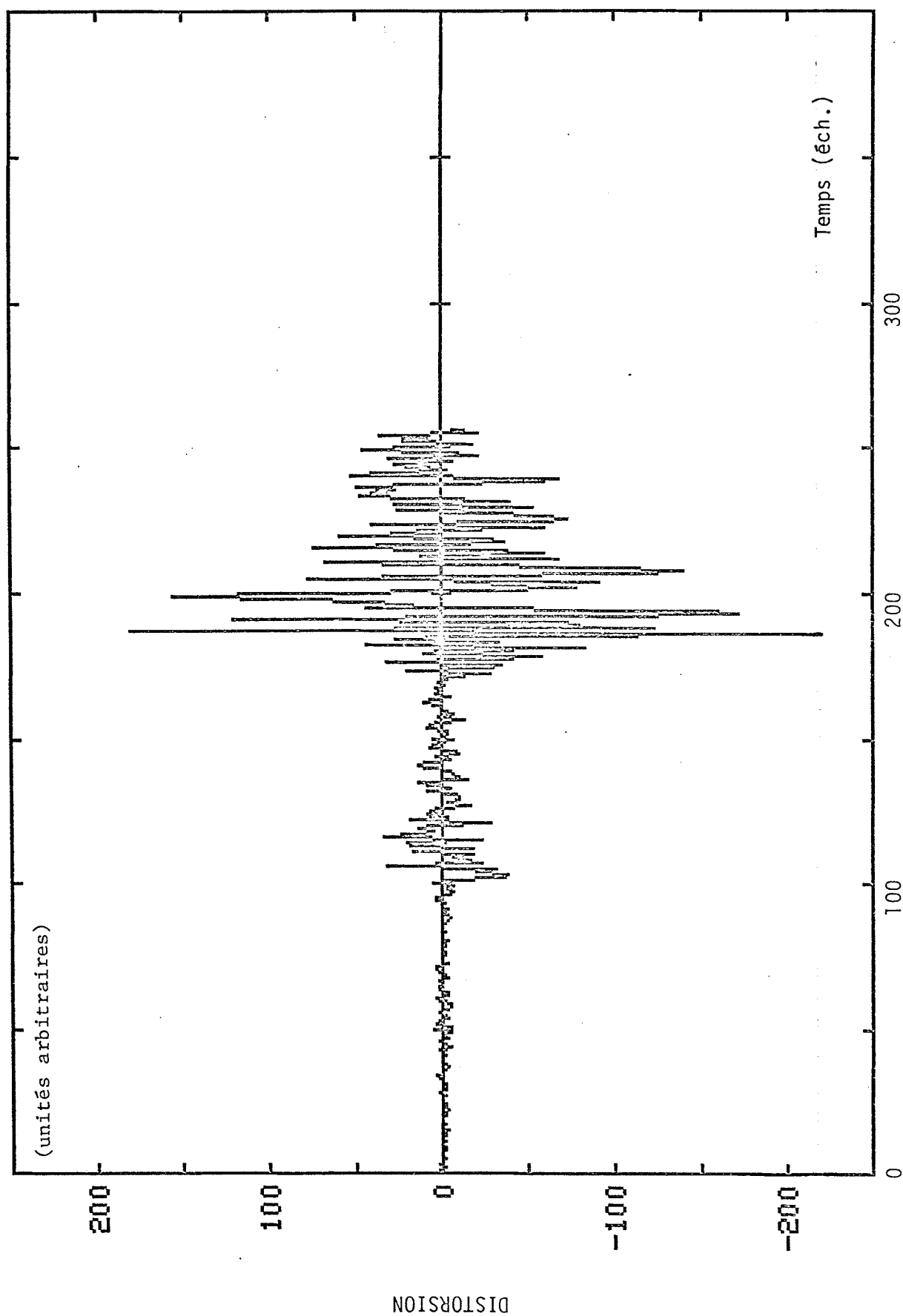


FIG 4-5 Distorsion en fonction du temps (allocations à taux fixe de 3 bits/éch.).

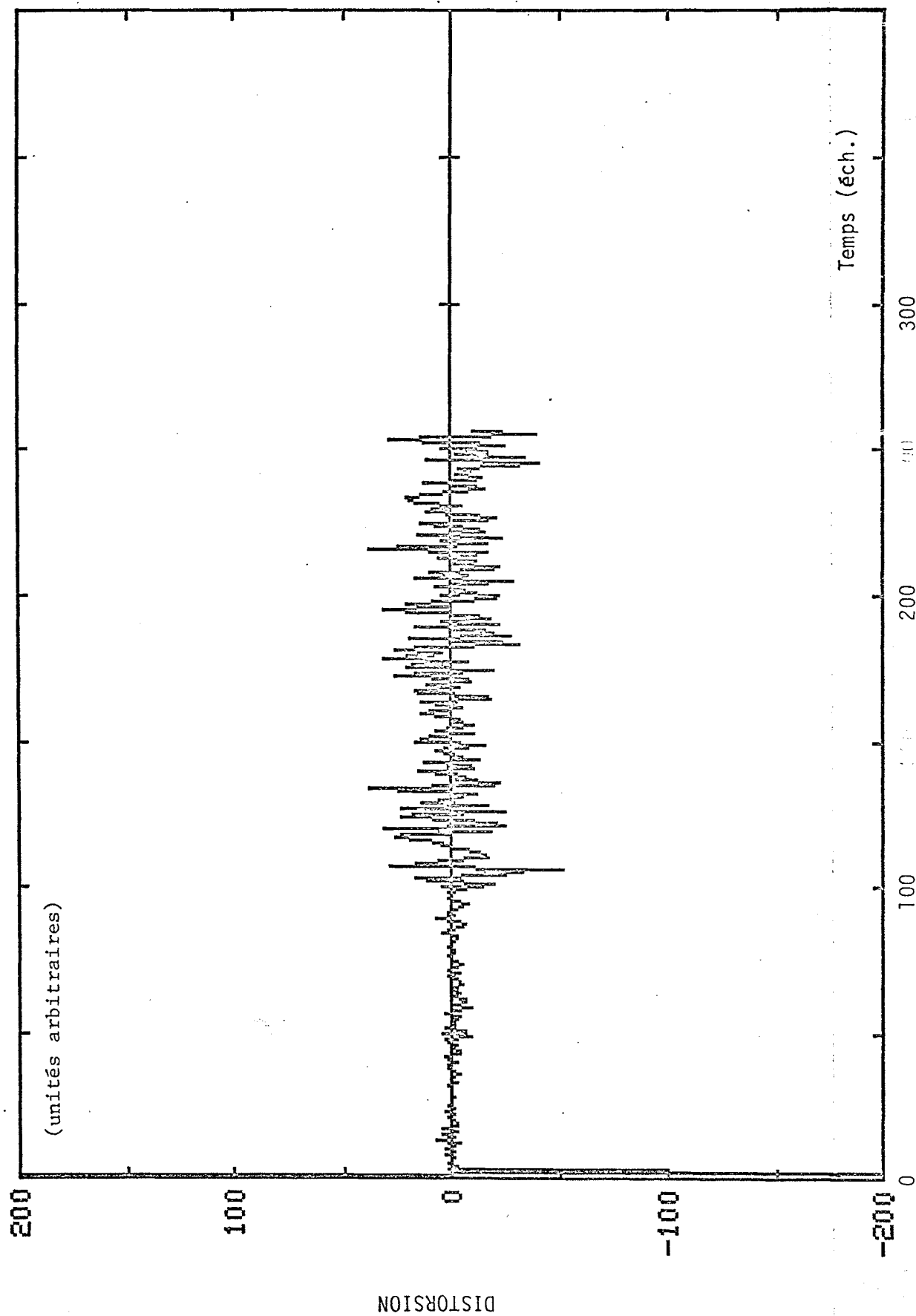


FIG 4-6 Distorsion en fonction du temps (allocations de la figure 4-4)

A la figure 4-5, on peut voir la distorsion en fonction du temps (erreur de quantification $Y(n)-X(n)$), pour le cas où le codeur fonctionne à taux fixe. A la figure 4-6, en revanche, on voit la même fonction d'erreur, mais cette fois pour le codeur fonctionnant à taux variable, selon les allocations montrées à la figure 4-4. En comparant les deux figures (4-5 et 4-6) on peut voir que la distorsion du codeur à taux variable est plus faible dans l'ensemble et qu'en plus, elle est effectivement plus également répartie sur la longueur du bloc, c'est-à-dire que la distorsion est plus faible (fig 4-6) là où elle était forte à taux fixe, et inversement, elle est plus forte là où la distorsion était à peu près nulle sur la figure 4-5.

Le codeur à taux variable a donc "rempli" sa mission qui est d'uniformiser la distorsion. Toutefois, on peut facilement constater, en regardant la figure 4-6 que la distorsion est encore loin d'être constante, comme le voudrait la théorie.

Dans le bloc particulier que nous étudions, les statistiques sont les suivantes:

Moyenne arithmétique des S_i^2	45194
Moyenne géométrique des S_i^2	2362
Gain théorique S/B std. (équ. 2.18).....	12.8 dB
Gain obtenu en pratique.....	9.4 dB
Variance de la distorsion (Taux fixe).....	251×10^5
Variance de la distortion (Taux var.).....	1×10^5

Les deux dernières statistiques nous indiquent donc que la distorsion carrée est beaucoup plus uniforme (sa variance étant environ 250 fois moindre à taux variable qu'à taux fixe), mais quand même loin d'être fixe (ce qui correspondrait à une distorsion ayant une variance nulle).

En somme, l'algorithme effectue bien son travail, mais ne peut naturellement pas obtenir une distorsion strictement constante, et ce, à cause des limitations physiques telles le fait que les allocations doivent se faire suivant un nombre entier de bits et que les plages d'allocations doivent être limitées avec une borne inférieure plus grande ou égale à zéro. L'hypothèse #2 est donc vérifiée comme étant la cause principale de la différence entre le gain théorique du S/B standard et les résultats obtenus par simulation.

Comme la formule de gain théorique (équ. 2.18) prend comme acquis que le gain obtenu est le maximum correspondant à une distortion fixe, il est de ce fait normal que cette formule soit considérée comme une

borne supérieure qu'on est loin de pouvoir atteindre en pratique.

Si on revient à la figure 4-1, on note que la courbe du S/B standard amorce une première montée rapide entre 0 et 32 éch., suivie d'un plateau entre 8 et 64 msec. et finalement une nouvelle montée par la suite. La première montée est attribuable aux variations de la fondamentale. En effet, nous avons vu au chapitre 2 que pour avoir un gain il faut qu'il y ait des variations importantes de la fonction de variance à l'intérieur du bloc considéré. Les variations importantes qui existent pour un bloc de 32 échantillons (4 msec.) se situent au niveau de la fondamentale. Par exemple, si on prend comme valeur typique $F_0 = 200$ Hz, on obtient une période de 4 msec.

Pour obtenir un gain supplémentaire, il faut attendre que la longueur du bloc dépasse la durée typique d'un phonème, soit environ 100 msec. (correspondant à un bloc de 1024 échantillons.). C'est en effet ce qu'on observe sur la courbe 4-1 pour le rapport S/B standard.

Le "plateau" que l'on remarque sur la courbe 2 de la figure 4-1 entre 8 et 32 msec. représente la transition entre ces deux types de variations. Ce plateau est mis en évidence du fait que l'allocation utilise un nombre entier de bits, ce qui demande une variation importante (min. rapport 2:1 entre d^2 et S_i^2) au niveau de l'enveloppe de la variance avant d'obtenir un gain supplémentaire. Ce n'est pas le

cas pour la formule de gain théorique (équ. 2.18) laquelle ne présente d'ailleurs pas de plateau.

On notera en outre que la courbe du gain du S/B standard "dépasse" le gain théorique (courbe 1) entre 3 et 10 msec..Ce dépassement n'est pas significatif puisqu'il s'agit au maximum d'une différence de 0.4 dB.

Si on considère maintenant la courbe 4-2a, on peut voir que le rapport S/B diminue à mesure que la constante de temps augmente, ce qui est normal. En effet, plus la constante de temps est longue, plus la fonction de variance s'éloigne de la forme du signal d'entrée rendant ainsi l'écart type de $P(XD_i | S_i^2)$ de plus en plus grande. Il se produit alors deux phénomènes:

- le premier, bien connu, est que l'adaptation du pas de quantification est beaucoup plus aléatoire. C'est pourquoi le rapport S/B diminue pour le codeur à taux fixe comme à taux variable.
- le second est que l'allocation devient erronée en ce sens que la fonction de variance varie trop lentement et n'est alors plus en mesure d'allouer d'une manière efficace. Voyons un exemple: Soit un ensemble X^2 pour lequel on calcule la fonction de variance avec une constante de temps courte ($W=2$): S_1^2 , et avec une constante de temps plus longue ($W=40$): S_1^2 . On prend $X_o^2=50$..

$$X^2 = \{ 2, 100, 4, 200, 3, 300 \}$$

$$S_1^2 = \{ 26, 63, 34, 117, 60, 180 \}$$

$$S_2^2 = \{ 49, 50, 49, 53, 51, 58 \}$$

Dans cet exemple, on voit que S_1^2 suit d'assez près les variations de X^2 quand $W=2$. Par contre pour $W=40$, la fonction ne suit absolument plus le signal d'entrée et accorde un poids très grand à la valeur accumulée ($\propto S_{i-1}^2$). Ceci a pour effet que $X_5^2=3$ recevra autant de bits que $X_1^2=100$. ou $X_4^2=200$..Donc, non seulement on ne fait pas mieux que le codeur à taux fixe (qui accorde 3 bits à chaque échantillon), mais on fait même pire encore, puisqu'on peut en arriver dans plusieurs cas à assigner plus de bits à des échantillons qui en ont un besoin moindre! C'est pourquoi on observe que le codeur à taux variable a une performance moindre que le codeur à taux fixe lorsque $W>10$.

4.2.3 Gain du rapport S/B segmentaire

A cause de sa nature, le rapport S/B segmentaire présente de plus grandes difficultés d'analyse que le rapport S/B standard.

Si on considère le rapport S/B segmentaire comme la moyenne géométrique des segments de 128 échantillons (réf. éq. 2.21), on

peut alors écrire:

$$4.02 \quad \left[\prod_{m=1}^M S/B_m^{(128)} \right]^{1/M} \leq \frac{1}{M} \sum_{m=1}^M S/B_m^{(128)} = \frac{1}{N} \sum_{i=1}^N S/B_i$$

où M est le nombre de segments sur un bloc de N échantillons, $S/B_m^{(128)}$ est le rapport S/B du segment m , et S/B_i est le rapport S/B du i ème échantillon (rapport en valeurs réelles et non en décibels.).

De l'équation ci-haut, il apparaît donc que le rapport S/B segmentaire, de par sa nature, devrait être théoriquement plus petit ou égal au rapport S/B standard. La relation d'égalité n'existe que si $S/B_m^{(128)}$ est constant pour tout m .

Nous avons tenté d'optimiser le rapport S/B segmentaire, d'une manière similaire à ce qu'on a fait au chapitre 2 pour le S/B standard. Ceci n'est toutefois pas possible parce que, dans le cas du S/B segmentaire, le différentiel dR/db_i (voir éq. 2.08) n'est pas indépendant des autres allocations $i \neq j$. En d'autres termes, on ne peut prétendre optimiser le rapport S/B segmentaire par un algorithme ou une relation simple entre b_i et S_i^2 comme c'est le cas pour le rapport S/B standard (éq. 2.14).

Si on considère l'équation B.07 de l'appendice B, que nous reproduisons ici:

$$4.03 \quad \left. \frac{S/B \text{ seg}}{I.V.} \right| = 20(C-K) \log 2 + 10 \log \frac{(\prod_{n=1}^M \bar{S}_n^2)^{1/M}}{(\prod_{i=1}^N S_i^2)^{1/N}}$$

où $\bar{S}_n^2$ est la variance moyenne du nième segment de 128 échantillons, on voit que tous les termes sont semblables à ceux de l'équation 2.17 pour le S/B standard, sauf pour le numérateur du terme d'extrême droite. Ce dernier, si on le compare à son équivalent de l'équation 2.17, à la lumière de l'équation 4.02, nous donne la preuve théorique que le gain du S/B segmentaire ne peut être plus grand que celui du S/B standard (pour plus de détails voir l'appendice B).

Maintenant, nous allons tenter d'expliquer le comportement du rapport S/B segmentaire sur la figure 4-1. Auparavant, nous introduisons la notion suivante:

La moyenne géométrique d'un processus quelconque diminue à mesure que son écart type augmente (Moyenne arithmétique tenue constante).

$$4.04 \quad \left[\prod_{i=1}^M R_i \right]^{1/M} < \left[\prod_{i=1}^M F_i \right]^{1/M} \text{ si } \sigma_R > \sigma_F \quad (\mu_R = \mu_F)$$

On sait que le codage à taux variable a pour but d'égaliser la distortion, ce qui fait que le rapport S/B d'un échantillon à l'autre devient très inégal. Si le bloc N sur lequel s'effectue l'assignement est plus petit ou égal à la longueur d'un segment (i.e. N 128), alors le rapport S/B segmentaire ne ressent pas ces fluctuations et

n'enregistre que la diminution de la distortion moyenne (Cette moyenne étant prise sur un segment de 128 échantillon, lequel selon notre hypothèse, est plus grand que le bloc utilisé). Ainsi, pour $N \leq 128$, la forme de la courbe de gain du S/B segmentaire devrait être sensiblement la même que celle du S/B standard (dans une proportion moindre toutefois, à cause des éq. 4.03 et 4.04). C'est effectivement ce qui se produit sur la figure 4-1 pour un délai variant entre 2 et 8 msec. alors que le rapport S/B segmentaire amorce une montée rapide comme le fait le S/B standard. Cette montée est due aux variations générées par les stries de la fondamentale et nous référons à la section précédente (S/B standard) pour le détail des explications.

Alors que le rapport S/B standard amorce une seconde montée à partir de $N=512$ (délai de 64 msec.), le rapport S/B segmentaire cette fois demeure stationnaire au même point, avec environ 2.8 dB de gain par rapport au taux fixe.

Dans le but d'expliquer cet "absence" de gain pour N grand, nous avons pris un exemple réel de rapports S/B pour une série de huit segments sur un bloc de 1024 échantillons:

Tableau VII - Effets de l'allocation variable sur l'écart type
des rapports S/B

Segment	Rapport S/B	
	Taux fixe	Taux variable
1	22.1	25.8
2	26.2	37.6
3	25.7	40.9
4	27.9	38.2
5	30.0	25.9
6	29.0	25.1
7	31.0	26.4
8	29.7	25.9
Moyenne des rapports:	27.7	30.7
Ecart type	: 2.73	6.4

L'écart type a donc plus que doublé. Ceci veut donc dire que même si la distorsion moyenne a diminué (augmentant ainsi le S/B standard), du point de vue segmentaire, elle a été "compensée" par une augmentation de l'écart type. Pour nous en convaincre, prenons un exemple numérique: soit un ensemble de rapports S/B $A_i = \{a_1, \dots, a_j\}$ provenant du codeur i avec rapport S/B a correspondant au segment j . Après codage, on suppose qu'on obtient les résultats suivants pour les algorithmes 1 et 2 respectivement:

$A_1 = \{ 2, 8, 20, 5, 50 \}$

$A_2 = \{ 4, 4, 4, 4, 332 \}$

Le rapport S/B standard, qui représente la moyenne arithmétique de A donne: $\bar{A}_1 = 17$. et $\bar{A}_2 = 69.6$, d'où un gain en faveur du codeur no. 2 (6.1 db). Si on fait maintenant la moyenne géométrique, on obtient: $\underline{A}_1 = 9.56$ et $\underline{A}_2 = 9.68$. Donc, d'un point de vue "segmentaire", le gain est absolument nul, alors que le gain S/B standard est de 6 db! Ceci est dû au fait que le coefficient de variation (écart type divisé par la moyenne arithmétique) est passé de 1.035 dans $\{A_1\}$ à 1.885 dans $\{A_2\}$, diminuant ainsi la moyenne géométrique d'une valeur équivalente au gain arithmétique (éq. 4.04).

Nos recherches nous ont convaincu que c'est exactement ce phénomène qui se produit avec le S/B segmentaire du codeur à taux variable. Le gain "arithmétique", que l'on retrouve lorsque $N \geq 512$, est dû aux variations de l'énergie par suite de la transition d'un phonème à l'autre. Ce gain est "compensé" par une "perte" due à l'augmentation du coefficient de variation chez les termes des rapports S/B d'un segment à l'autre.

Concernant le comportement du rapport S/B segmentaire en fonction de la constante de temps de la variance (fig 4-2), l'effet est absolument le même qu'avec le S/B standard. L'explication donnée à

la section 4.2.2 vaut donc aussi pour le rapport S/B segmentaire.

Ce qu'on retient avant tout, c'est le comportement à priori surprenant du rapport S/B segmentaire, lequel nous dicte une limite supérieure au gain du codage à taux variable et cette limite se situe en-deça de 3 dB.

4.3 JUGEMENTS SUBJECTIFS

Avant d'aller plus avant avec des modifications possibles de l'algorithme, il convient de vérifier si l'amélioration apportée par le codeur à taux variable en vaut la peine. La méthode la plus efficace et universellement reconnue est de procéder à une série de tests subjectifs.

4.3.1 Méthode utilisée

Il existe plusieurs méthodes pour effectuer les tests subjectifs [28]. Celui que nous avons utilisé, et qui est le plus courant, est la procédure d'isopréférence [29]. Pour être bref, disons que le test consiste à comparer des phrases codées par différents algorithmes avec les mêmes phrases codées par un algorithme connu (dans notre cas

"Log-PCM") utilisant différents taux de transmission (4,5,6 ou 7 bits/échantillon).

Chaque test consiste en une séquence 'ABAB' et le sujet doit déterminer laquelle de A ou B il préfère. Les résultats sont alors compilés sur ordinateur dans le but de déterminer une préférence "équivalente" entre les phrases test et l'échelle de mesure (ex: 6.5 bits log-PCM).

Dans le cas qui nous concerne, les mesures obtenues sont le résultat de 48 tests effectués par dix sujets dont la moitié pourrait être qualifiée de "sujets inexpérimentés". Deux phrases, l'une française, l'autre anglaise, ont été utilisées.

4.3.2 Résultats des tests

Nous donnons dans le tableau ci-après le codeur ayant fait l'objet de test ainsi que le résultat de la projection d'isopréférence sur une échelle "log-PCM".

Le codeur appelé "ADPCM-F" correspond à un codeur différentiel PCM avec quantificateur adapté suivant les multiplicateurs de Jayant et

utilisant un prédicteur fixe [8,10]. Le codeur "OPTA" correspond à l'algorithme que nous avons conçu (c.f. chap III) et utilisant les paramètres optimaux suivants:

Longueur de bloc.....256 échantillons

Constante de temps..... 2 échantillons

Plage d'allocation.....2 à 6 bits (taux var.)

3 bits (taux fixe)

Le tableau qui suit donne les résultats des tests:

Tableau VIII - Résultats des tests subjectifs

Codeur	Taux utilisé (bit/éch.)	Taux équivalent (Log-PCM)
1) ADPCM-F	3	6.10
2) ADPCM-F	4	6.92
3) OPTA (T.FIXE)	3	6.47
4) OPTA (T. VAR.)	3	6.54

A la lumière du tableau, on peut d'abord tirer la conclusion que le codeur 'OPTA' se situe, en termes de qualité subjective, à mi-chemin entre le ADPCM standard à 24 Kbits/sec. et le même codeur fonctionnant à un taux de 32 Kbits/sec..

Deuxièmement, et c'est cela qui frappe encore plus, le codage à taux variable n'apporte pratiquement rien de plus. En fait, tout le gain provient du délai (ou ajustement du pas de quantification "en amont"), tel que nous l'expliquons à la section 4.1.2 .

Il est effectivement difficile d'apprécier la différence entre une phrase codée par l'algorithme 'OPTA' que ce soit à taux fixe ou à taux variable. On peut invoquer pour expliquer un si faible écart subjectif que, suivant le rapport S/B segmentaire, il y a une différence de moins de trois (3) décibels entre les deux méthodes (27.5 dB à taux variable contre 24.7 à taux fixe). On s'accorde généralement pour dire que la marge de trois décibels représente le seuil en-deça duquel l'oreille humaine peut difficilement apprécier une différence de qualité.

Nous avons aussi comparé jusqu'à quel point les mesures objectives de bruit s'accordaient avec la mesure subjective. Pour cela, nous avons calculé les mesures objectives (standard et segmentaire) des phrases "étalon" log-PCM et avons ainsi obtenu par interpolation, une

mesure de S/B subjective pour les phrases tests.

L'erreur moyenne (RMS) entre les mesures objectives et subjectives, dans le cas du rapport S/B standard, est de 5.97 dB, alors qu'elle diminue à 2.92 dB dans le cas du rapport segmentaire. Ce résultat peut être comparé aux mesures de M. Nakatsui [33], qui obtient une erreur de 7.31 dB pour la mesure standard contre 4.78 dB pour la mesure segmentaire.

Du point de vue des mesures objectives de bruit, le rapport S/B segmentaire est donc plus pertinent, et c'est donc l'optimisation de ce dernier qui nous intéressera plus particulièrement dans la prochaine section.

4.4 OPTIMISATION DU RAPPORT S/B SEGMENTAIRE

Nous avons vu, à la section 4.2.3, que l'algorithme de codage à taux variable agit sur le rapport S/B segmentaire suivant deux tendances antagonistes: la première vise à augmenter ce rapport par la diminution de la distorsion moyenne, tandis que la seconde vise à diminuer ce même rapport par suite de l'augmentation du coefficient de variation des rapports S/B de chaque segment. Nous avons pu constater qu'en "jouant" sur ces deux tendances (notamment dans la figure 4-1 et

le tableau V), on obtient un maximum correspondant aux paramètres optimaux mentionnés à la section 4.3.2:

- a) Capacité de mémoire tampon d'entrée.....256 éch.
- b) Constante de temps (Variance)..... 2 éch.
- c) Plage d'allocation2 à 6 bits

On note ci-haut que les caractéristiques "a" et "c" visent justement à limiter le coefficient de variation. En effet, si par exemple, on permet une plage de 0-6 bits, les variations du S/B d'un segment à l'autre seront plus fortes que si on limite le minimum à deux bits.

Les figures ne nous assurent cependant pas de l'optimalité du S/B segmentaire (nous ne pouvons exprimer le maximum du rapport S/B segmentaire de façon analytique et c'est là tout le problème: nous devons espérer que les essais sont "exhaustifs" avant de déclarer que le maximum possible de gain du rapport S/B segmentaire est en moyenne moindre que trois décibels...). C'est pour ces raisons que nous avons décidé de tenter d'autres approches propres à optimiser le rapport S/B segmentaire. En fait, ce que nous cherchons, c'est d'avoir un bloc d'échantillons codé avec l'algorithme standard (c'est en effet le

moyen le plus efficace pour diminuer la distorsion moyenne carrée) tout en ayant un coefficient de variation des segments qui soit moindre par rapport à ce que nous obtenons jusqu'ici.

Nous avons vu que la fonction de S/B segmentaire est telle que, contrairement au S/B standard, on ne peut opérer une allocation sur un échantillon donné sans influencer le reste des échantillons. Nous avons donc imaginé, pour des fins purement "académiques", de faire une allocation en deux phases, la première suivant la méthode de gain marginal habituelle, et la seconde dans le but d'ajuster les allocations à la lumière des effets obtenus lors de la première allocation.

Suivant cette philosophie, nous avons essayé plusieurs algorithmes qui ont tous donné des résultats moindres ou équivalents au maximum déjà obtenu sur les figures de ce chapitre. Nous vous donnons à titre d'exemple les quatre étapes de la troisième et dernière version d'allocation "itérative" que nous avons essayé :

- 1) Codage d'un bloc de 1024 échantillons par l'algorithme standard.
- 2) Mesure du rapport S/B sur chaque segment du bloc (huit en tout) et calcul de la moyenne.
- 3) Calcul de la différence en décibels entre un segment et la moyenne du bloc. Assignment de bits supplémentaires en fonction de cette différence (à raison de 1 bit/éch pour chaque 6 dB de

différence).

4) Recodage à taux variable sur chaque segment de 128 échantillons (et non plus 1024 comme en "1") avec un nombre de bits tel qu'ajusté en "3".

Par cet algorithme, nous espérons tirer les avantages du codage à taux variable, tout en évitant un accroissement trop important de l'écart type des rapports S/B sur chaque segment.

Par exemple, si après la première "ronde" de codage, un segment totalise R bits et possède un rapport S/B de 3 dB supérieur à la moyenne, il sera recodé séparément à la deuxième "ronde", mais cette fois avec $(R - 3 \times 128 / 6)$ bits, soit 64 bits de moins.

Le résultat obtenu est de 24.5 dB pour le rapport S/B standard et de 28.2 dB pour le rapport S/B segmentaire. Avec le même fichier codé suivant la méthode standard (bloc de 1024 échantillons), on obtient 29.2 dB et 28.7 dB respectivement. Nous n'avons donc réussi qu'à obtenir un résultat équivalent (du point de vue S/B segmentaire) à l'algorithme standard d'allocation.

L'allocation en deux phases, telle que proposée, ne permet donc pas d'accroître le rapport S/B segmentaire au-delà de ce qu'on avait

déjà obtenu avec l'allocation en une seule phase. Il semble bien que la théorie proposée au chapitre II pour maximiser le S/B standard soit adéquate pour maximiser aussi le rapport S/B segmentaire.

4.5 DEPENDANCE DU TAUX DE CANAL

On peut se poser une dernière question sur les résultats obtenus jusqu'ici: est-ce que le comportement du gain de codage à taux variable demeure le même lorsque le débit moyen de canal change?

Nous avons donc effectué quelques essais avec un débit de canal de 1 bit par échantillon (8000 bits par seconde) avec les résultats que voici (constante de temps de la variance: 2 échantillons; plage d'allocation: 0 @ 3 bits):

Tampon d'entrée	S/B std.	S/B seg.
(Taux fixe)	9.7	12.2
256	7.7	9.9
1024	10.0	9.5

Ainsi donc, l'effet du codage à taux variable est encore pire si le débit de canal est très bas. Ceci est probablement dû au fait qu'à un

débit de 1 bit par échantillon, la différence de qualité entre une allocation nulle et une autre de 2 bit est très grande, ce qui affecte encore plus le S/B segmentaire. En effet, le coefficient de variation augmente d'un facteur de plus de 17 (0.065 à taux fixe contre 1.12 à taux variable) alors que l'augmentation n'est que du double lorsque le taux de canal est de trois bits par échantillon (voir section 4.2.3). Les jugements subjectifs confirment d'ailleurs les mesures objectives.

V.- ESSAIS DE REALISATIONS PRATIQUES

5.1 - INTRODUCTION

Au cours du dernier chapitre, nous avons vu que le codeur à taux variable, placé dans des conditions idéales (i.e. sans tenir compte du délai ou du taux d'information latérale), nous donne un gain de moins de 3 décibels (rapport S/B segmentaire).

Devant une telle situation, il apparaît téméraire de vouloir aller plus loin et construire une version pratiquement réalisable à partir de l'algorithme du chapitre IV ou un autre. En effet, dans notre effort de limiter le taux d'information latérale (i.e. la transmission de la fonction de variance ou l'équivalent) à un pourcentage "raisonnable" par rapport à l'information primaire (ex.: 10%), il faudra songer à une forme de variance qui soit de type syllabique et non instantanée. Or, nous avons vu que l'utilisation de la variance instantanée semble être une condition "sine qua non" pour obtenir un gain par rapport au codage à taux fixe (c.f. figure 4-2). En conclusion, quelle que soit notre choix pour le calcul et la transmission de la variance (ou encore du pas de quantification et de l'allocation), on ne peut s'attendre à beaucoup mieux que le même codeur fonctionnant à taux fixe.

Toutefois, pour être complet, nous avons voulu modifier l'algorithme "OPTA" (chap III et IV) pour obtenir une version réalisable et en étudier le comportement. En outre, dans la dernière section du présent chapitre, nous apporterons certains éclaircissements sur l'article de J.J. Dubnowski et R.E. Crochière concernant le codage à taux variable [23] ainsi que les essais que nous avons faits sur l'algorithme.

5.2 CODEUR "OPTA" MODIFIE

Dans le but de rendre l'algorithme du codeur "OPTA" réalisable en pratique, nous devons nous attaquer à trois problèmes: la capacité de mémoire (délai), la transmission d'information latérale et finalement le temps d'exécution des calculs.

5.2.1 Limitation sur la capacité des mémoires tampon

La capacité des mémoires tampon est directement proportionnelle au délai encouru durant la transmission. Si on regarde la figure 5-1, on peut voir qu'il existe trois types de mémoires:

a) tampon d'entrée pour la transmission: Utilisé dans le but d'accumuler la valeur d'un bloc de N échantillons.

b) Les deux autres tampons fonctionnent en paires. Par exemple, si

la mémoire sortie du transmetteur est à 75% remplie, la mémoire d'entrée du récepteur sera alors à 25% de sa capacité. La capacité totale est donc égale à la capacité de l'une ou l'autre des mémoires tampon, chacune ayant une capacité égale.

A l'équilibre, il faut que les tampons soient à 50% remplis à la fin d'un bloc donné. A ce moment, le pire qui puisse se produire est que les $N/2$ prochains échantillons soient codés avec 0 bits, et les $N/2$ derniers codés avec $2C$ bits (C étant le taux de canal). Dans la première moitié du bloc, le tampon de sortie se vide d'une valeur $NC/2$ bits et dans la seconde moitié, il se remplit d'un montant équivalent, retournant ainsi à son point de départ (50%). La capacité minimum qu'on doit fournir est donc NC bits, ou N échantillons.

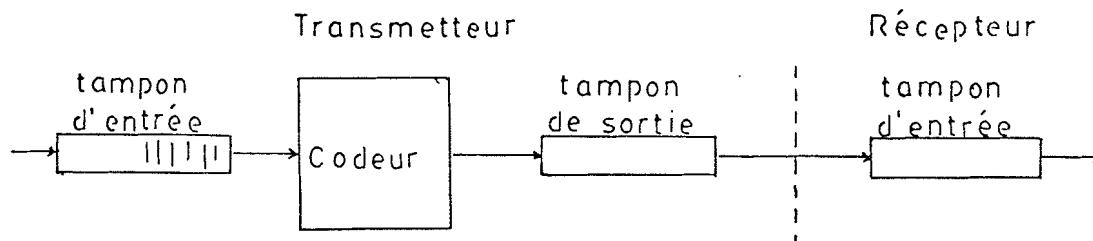


FIG 5-1 Agencement des mémoires dans le codeur "OPTA" modifié.

Le délai total est donc de $2N$ pour ce type de codeur (N pour le premier tampon d'entrée, et N pour les deux autres mémoires fonctionnant en paires). Comme nous l'avons vu au chapitre précédent, l'optimum est atteint lorsque N est égal à 256 échantillons. Avec un tel bloc, le délai total atteint alors 512 échantillons ou 64 millisecondes (ou délai aller-retour de 128 msec.).

Dans certaines conditions d'opération, un tel délai pourrait forcer l'utilisation d'un supprimeur d'écho (la limite suivant les normes nord-américaines étant 45 msec.).

Le problème du délai en lui-même n'est pas trop grave lorsqu'on songe que celui des liens satellites est de l'ordre de 500 millisecondes. On peut consulter à ce sujet les études subjectives réalisées par G. K. Helder en 1964-65 [36].

5.2.2 Limitation de l'information latérale

Dans ce cas-ci, à cause du fait que l'estimation du pas et l'allocation sont effectuées avant même la quantification ("forward estimate"), il est nécessaire de transmettre soit la fonction de variance (à partir de laquelle le récepteur reconstitue la mise à jour du pas et de l'allocation) soit les valeurs du pas de quantification et de l'allocation. Nous avons choisi cette dernière solution puisqu'elle diminue considérablement la complexité du récepteur et augmente son immunité aux erreurs de transmission.

Puisque nous ne pouvons transmettre qu'un nombre limité d'information de ce type (On fixe le maximum à environ 2400 bits/secondes) il va de soi que le pas et l'allocation ne doivent pas

changer à chaque échantillon comme c'était le cas au chapitre IV.

Avec 3 bits, on peut coder toute les allocations possibles (de 0 à 6 bits) et ce, sans erreur puisque les allocations sont entières. Concernant la quantification de la valeur du pas des quantificateurs, nous avons opté pour un codage à quatre bits utilisant une échelle logarithmique ($\mu = 255$).

Les valeurs de mise à jour peuvent être transmises tous les m échantillons. Pour ce faire, on peut utiliser une variance syllabique, ou encore une variance instantanée dont on fait une moyenne sur m échantillons. C'est cette dernière approche que nous avons retenue puisqu'elle donne des résultats légèrement meilleurs. Dépendant de la valeur de m , on obtient différents taux d'information latérale:

m	Taux d'information	% du taux primaire
8	7000	29.0%
16	3500	14.6%
32	1750	7.3%
64	875	3.6%
128	437.5	1.8%

Ainsi, pour $m=64$, cela signifie qu'on met à jour l'allocation et le pas de quantification tous les 64 échantillons, soit 4 fois par bloc de 256. Du tableau ci-haut, il apparaît qu'on ne peut aller au-dessous de $m=32$, puisqu'alors le taux devient trop important.

5.2.3 Temps d'exécution des calculs

Le calcul de la variance ne présente pas de problème puisqu'on peut l'exécuter durant le remplissage du tampon d'entrée. Toutefois, les allocations doivent être calculées une fois terminée l'accumulation des N échantillons. Pour cela, on peut augmenter la capacité de la mémoire tampon. Disons par exemple qu'on dispose d'un temps supplémentaire de 32 ms. (256 éch.) pour effectuer les allocations. Pour exécuter le calcul pour huit allocations ($m=32$), il faut d'après l'équation C.06 (appendice C): 216 opérations. Chaque opération doit donc durer au plus 150 nanosecondes, ce qui est réalisable.

5.2.4 Description globale de l'algorithme OPTA modifié

A partir des contraintes mentionnées dans les sous-sections précédentes, on modifie donc le codeur OPTA du chapitre III selon les données suivantes:

1) Variance

On calcule une fonction de variance instantanée ($\alpha=0.5$). Ensuite, on calcule une valeur moyenne de cette fonction pour des "sous-blocs" de m échantillons:

$$5.01 \quad \overline{S}_n^2 = \frac{1}{m} \sum_{i=(n-1)m}^{nm-1} S_i^2$$

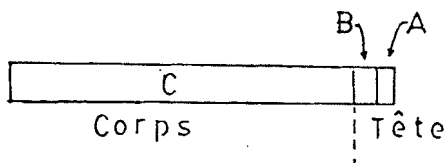
2) Allocation et pas de quantification

On évalue ces vecteurs de la même façon qu'au chapitre III, sauf qu'on utilise pour ce faire la fonction de variance $\overline{S}_n^2$ au lieu de S_i^2 , i.e. on calcule N/m différentes valeurs de Δ_i et b_i au lieu de N .

3) Transmission de l'information latérale

L'allocation des bits est transmise selon un code de 3 bits et l'ajustement du pas selon un code de 4 bits log-PCM ($\mu=255$), pour un total de 7 bits pour chaque mise à jour. On pourrait par exemple imaginer que ces bits soient intercalés au début de chaque bloc.

Ci-dessous, un exemple avec $N=256$ et $m=32$:



A.....24 bits pour les allocations (8x3 bits)

B.....32 bits pour l'ajustement du pas
de quantification (8x4 bits)

C.....768 bits de données (256x3 bits)

Au total, on a 824 bits à être transmises dans un délai de 32 msec. (256 éch. à 0.125 msec/éch.) soit un taux net de 25,750 bits par seconde.

5.2.5 Résultats de l'algorithme OPTA modifié

Si on réfère à la figure 5-3, on remarque que le rapport S/B diminue uniformément à mesure que la longueur du sous-bloc augmente, ce qui est tout-à-fait normal puisque la mise à jour du pas et de l'allocation se fait à un rythme de plus en plus lent à mesure que m augmente.

Pour fins de référence, nous avons inclus le même codeur

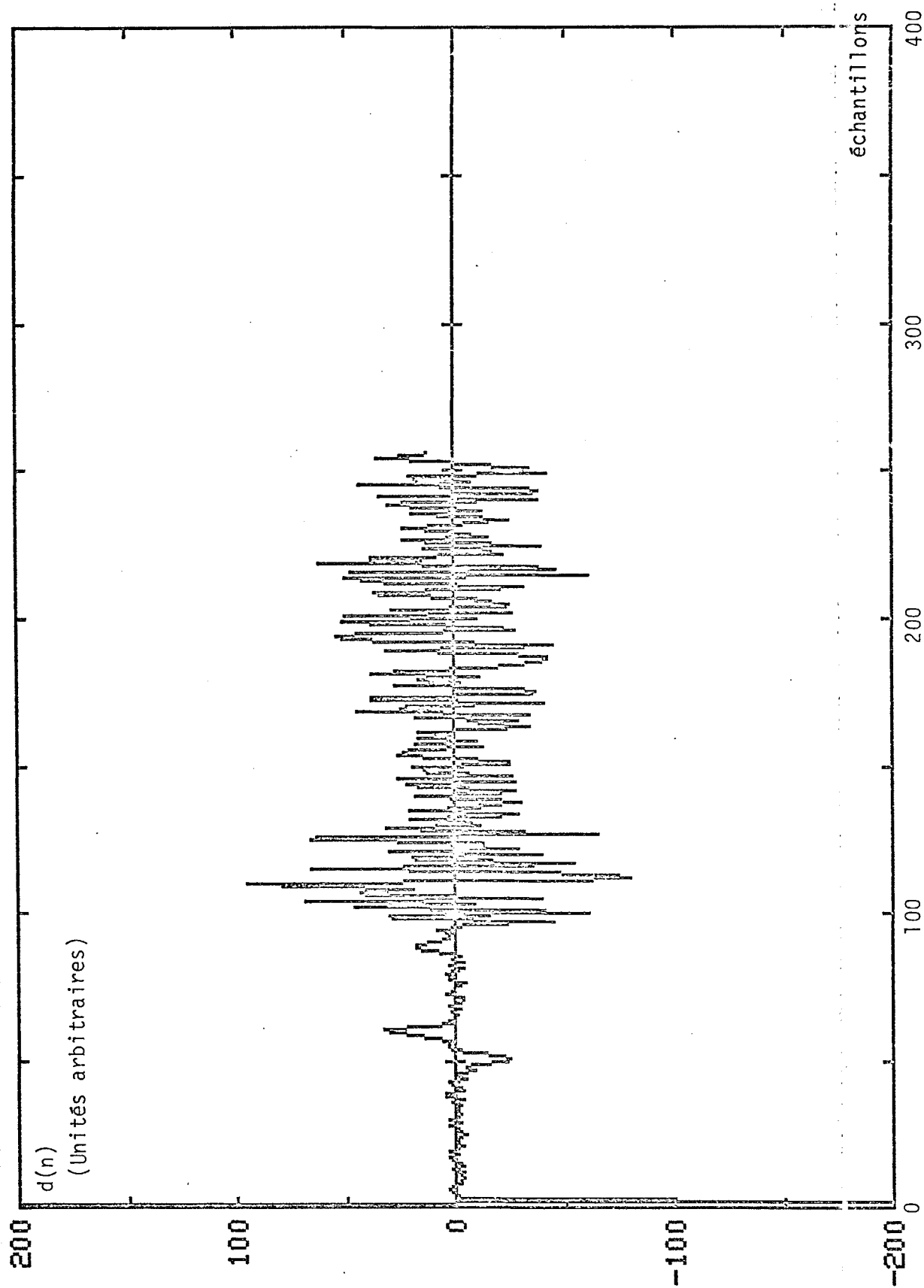


FIG 5-2 Distorsion d'un bloc du codeur "OPTA" modifié correspondant aux figures 4-5 et 4-6 du codeur OPTA idéal (chap IV).

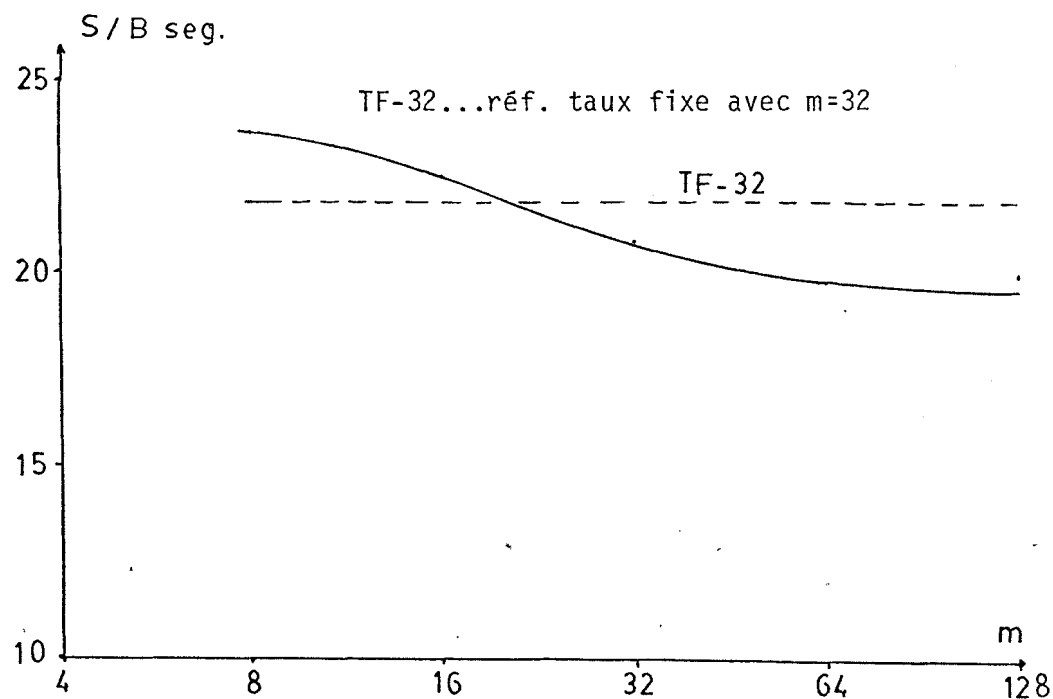
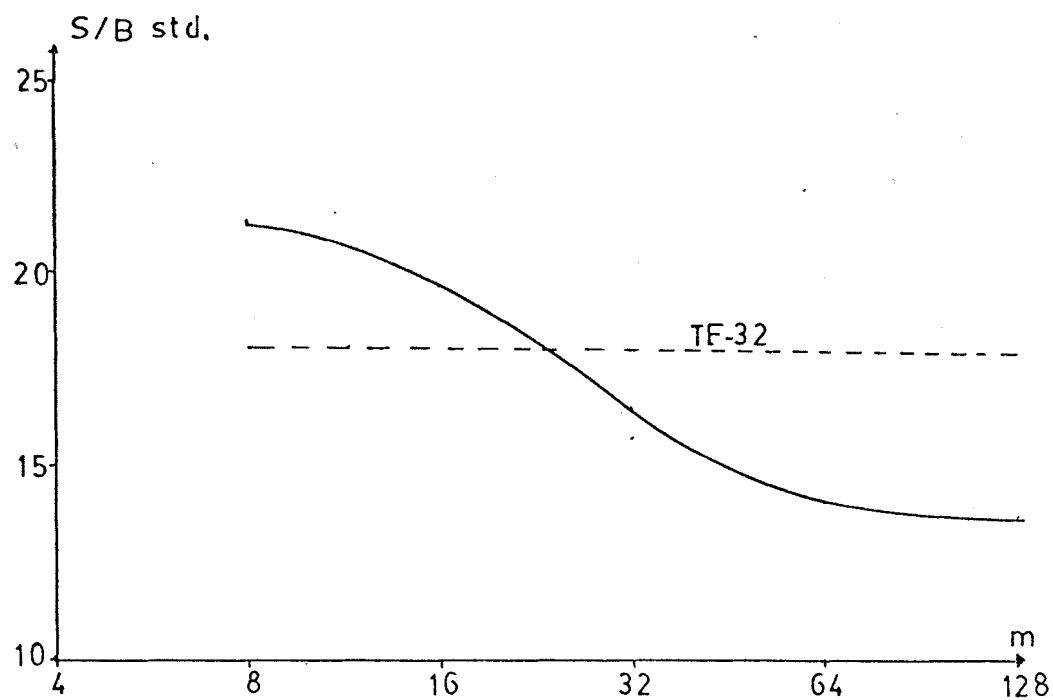


FIG 5-3 Comportement du rapport S/B en fonction de la fréquence des mises à jour (m) du pas de quantification et de l'allocation - codeur OPTA modifié.

fonctionnant à taux fixe avec $m=32$ (noté T.F.-32). On peut alors voir que le taux fixe est supérieur au taux variable avec le même sous-bloc de 32 échantillons. Suivant l'analyse théorique que nous avons élaboré au chapitre précédent, ainsi que d'après la figure 4-2, les résultats obtenus ne doivent pas nous surprendre.

En fait, le gain du codage à taux variable n'atteint que 1.1 dB pour des sous-blocs de huit échantillons, ces derniers nécessitant 7000 bits par seconde en terme d'information latérale...

Il n'existe donc aucun moyen de modifier l'algorithme pour obtenir un gain minimum attribuable à l'allocation variable, tout en limitant le taux d'information latérale.

On peut dire, en contrepartie, que le même algorithme fonctionnant à taux fixe et utilisant des sous-blocs de 32 échantillons nous donne un rapport S/B segmentaire de 21.8 dB (25 Kbs) alors que l'algorithme ADPCM standard (avec multiplicateurs de Jayant) donne un S/B de 19.0 dB, soit un gain de 2.8 dB attribuable au délai. Lorsqu'on mesure subjectivement les deux méthodes, il est cependant difficile d'apprécier cette amélioration.

5.3 ALGORITHME DE DUBNOWSKI ET CROCHIERE

5.3.1 Introduction

Dans la revue "BSTJ" de mars 1979 [23] apparaissait un article intitulé: "Variable Rate Coding of Speech". Cet article a été le point de départ de tout ce travail.

Selon les auteurs, le codage à taux variable ne devient intéressant que pour un délai de 100 millisecondes ou plus, ce qui, toujours selon les auteurs, est très élevé en pratique. Pour cette raison, ceux-ci ne recommandent l'usage de cet algorithme qu'avec plusieurs usagers en parallèle (ex.:T.A.S.I.).

Le but de la présente section est donc de "compléter" l'article de Dubnowski et Crochière en présentant les mesures de rapports signal/bruit qu'on obtient avec leur algorithme.

5.3.2 Description de l'algorithme

Notre but n'est pas de faire une description détaillée de

l'algorithme, celle-ci étant déjà disponible à la référence [23] de la bibliographie.

Si on se réfère à la figure 5-4, on trouvera le schéma bloc de l'algorithme. Ce dernier consiste en une mémoire tampon d'entrée dans laquelle on vient puiser des échantillons à un rythme variable, suivant l'allocation en cours.

L'allocation elle-même est effectuée suivant l'équation 2.14. Toutefois, la variance est dérivée du pas de quantification [27]. Quant à la distorsion d^{*2} , on part avec une valeur initiale quelconque et on l'ajuste à la hausse si le tampon se remplit au-delà de 50% de sa capacité, et à la baisse dans le cas contraire.

Dans le but de limiter l'information latérale, on calcule une allocation à un moment donné et on la conserve ainsi jusqu'à ce que l'on ait formé un paquet de 60 bits (Si l'allocation est de 3 bits, on code ainsi vingt échantillons dans un paquet, alors qu'avec six bits, on n'en code que dix; c'est pourquoi le taux d'occupation du tampon d'entrée varie avec l'allocation). La valeur de l'allocation est codée (2 ou 3 bits) et forme la tête du paquet.

Il en résulte donc un système dont le seul délai est la mémoire

tampon d'entrée et dont la sortie consiste en paquets émis à intervalles réguliers (taux de 25200 bits par secondes pour une précision moyenne de 3 bits/éch. et 3 bits pour l'allocation).

5.3.3 Résultats obtenus

Nous présentons les résultats sous la forme de trois tableaux. Avant de les consulter, il serait bon de savoir que le même fichier audio, codé selon cet algorithme fonctionnant à taux fixe, a donné le résultat suivant:

S/B standard.....14.4 dB.

S/B segmentaire.....20.6 dB.

Tableau IX : Comportement en fonction de la capacité
de la mémoire tampon d'entrée.

Capacité	S/B std.	S/B seg.
256	14.4 dB	19.4 dB
512	15.2	19.2
1024	15.6	18.7

Tableau X : Comportement en fonction de la longueur du paquet

Longueur*	S/B std.	S/B seg.
30	14.4 dB	20.0 dB
60	14.4	19.4
90	14.6	19.2
120	15.0	19.3

* Nous avons dû empêcher les allocations de 4 bits puisque 30 et 90 ne sont pas multiples de quatre.

Tableau XI: Comportement en fonction du multiplicateur $H(b(n-1))$

(Voir figure 5-4 ou réf. [23])

$H(b(n-1))$	S/B std.	S/B seg.
1.001	14.3 dB	19.8 dB
1.01	14.4	19.4
1.05	14.2	19.3
1.10	13.6	19.7
1.2	13.7	19.4

5.3.4 Commentaires sur les résultats

Même si, à prime abord, l'algorithme semblait très intéressant, on est à même de constater que les résultats des mesures segmentaires

sont moindres qu'à taux fixe.

Comme l'article original est basé sur un taux de canal de 32 Kbits/sec., nous avons aussi fait des essais avec quatre bits par échantillon. Dans le meilleur des cas, nous obtenions un gain (S/B standard) de 3.8 dB, cependant que le gain S/B segmentaire n'était que 0.5 dB.

A la lumière des études que nous avons effectué jusqu'ici, nous pouvons dire que le problème vient du fait que les ajustements sont d'ordre "syllabiques". Nous avons vu au chapitre IV que les mises à jour des allocations doivent être instantanées si on veut avoir un minimum de gain. Si on se réfère à la figure 4-2, c'est comme si on se situait à $W=12$, là où on a une perte du point de vue segmentaire, et un gain du point de vue S/B standard.

A la page suivante, on montre les valeurs d'allocation (fig 5-5), la fonction de variance (fig 5-6) et le comportement du tampon d'entrée (fig 5-7) en fonction des paquets de sortie. En superposant les figures 5-5 et 5-6, on peut voir que les allocations suivent les variations de l'enveloppe de la fonction de variance dans la mesure où le tampon d'entrée n'est pas saturé (fig 5-7).

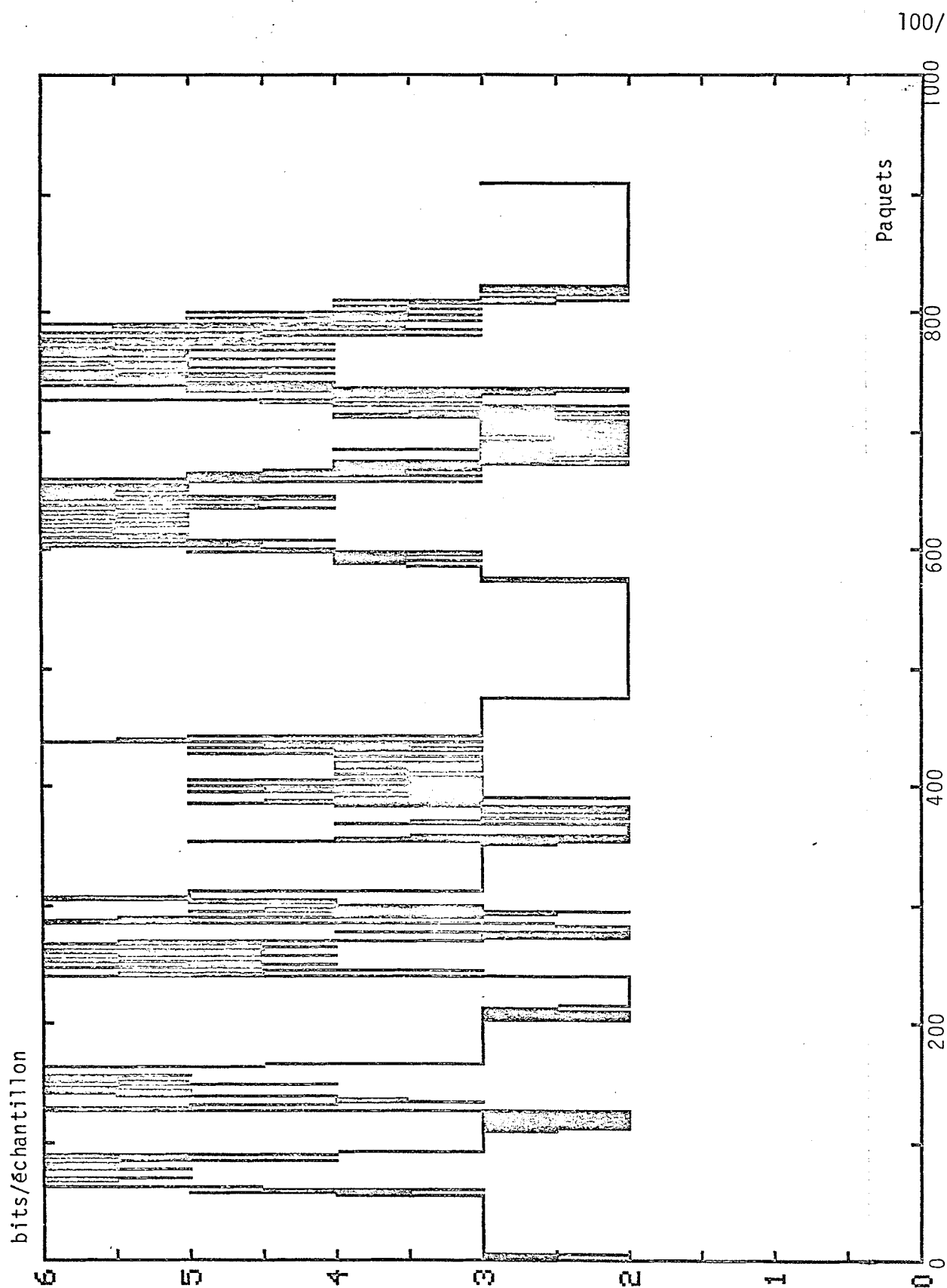


FIG 5-5 Allocations en fonction du temps (tampon de 1024 éch. et $H(b(n))=1.01$)

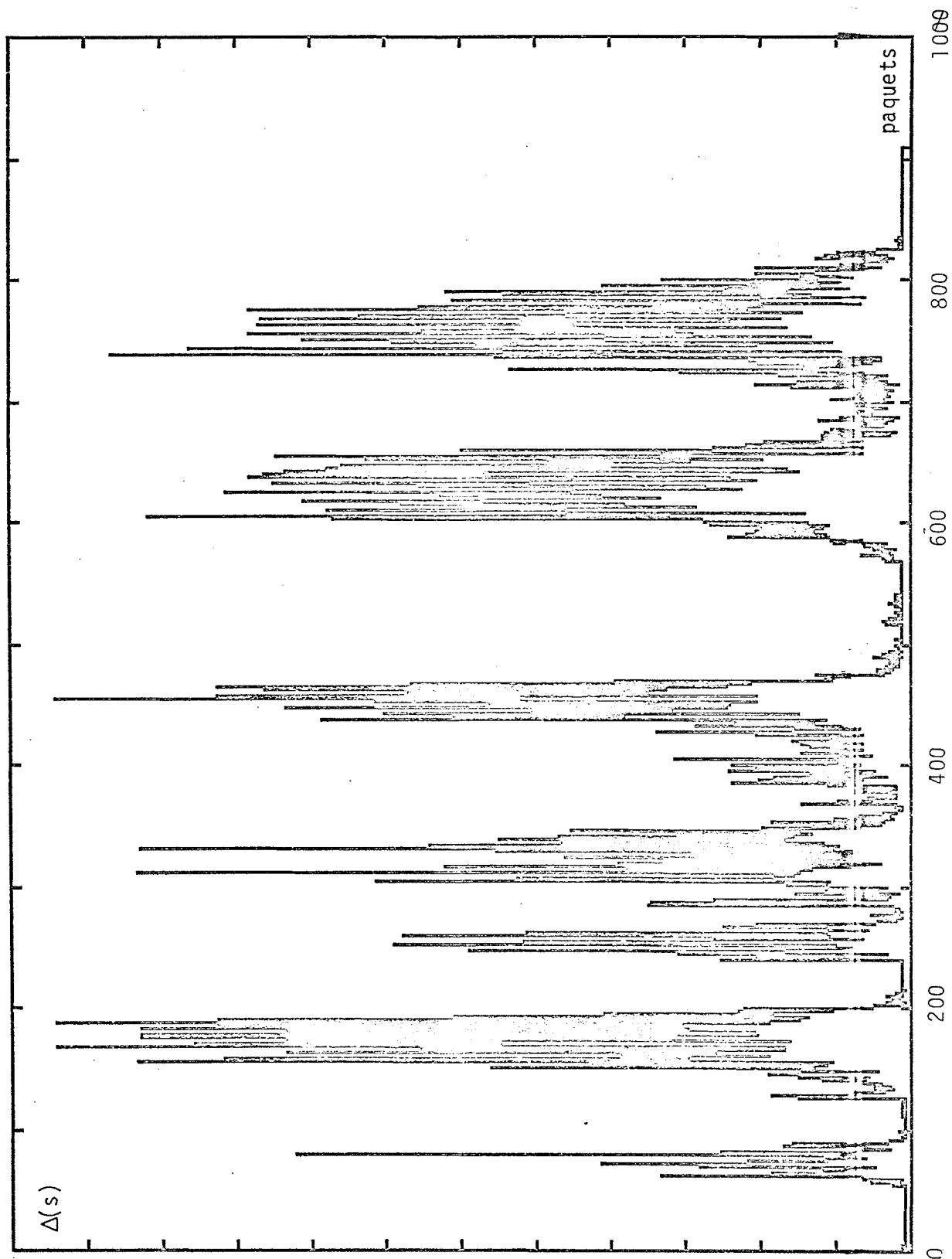


FIG 5-6 Variations du pas de quantification d'un paquet à l'autre

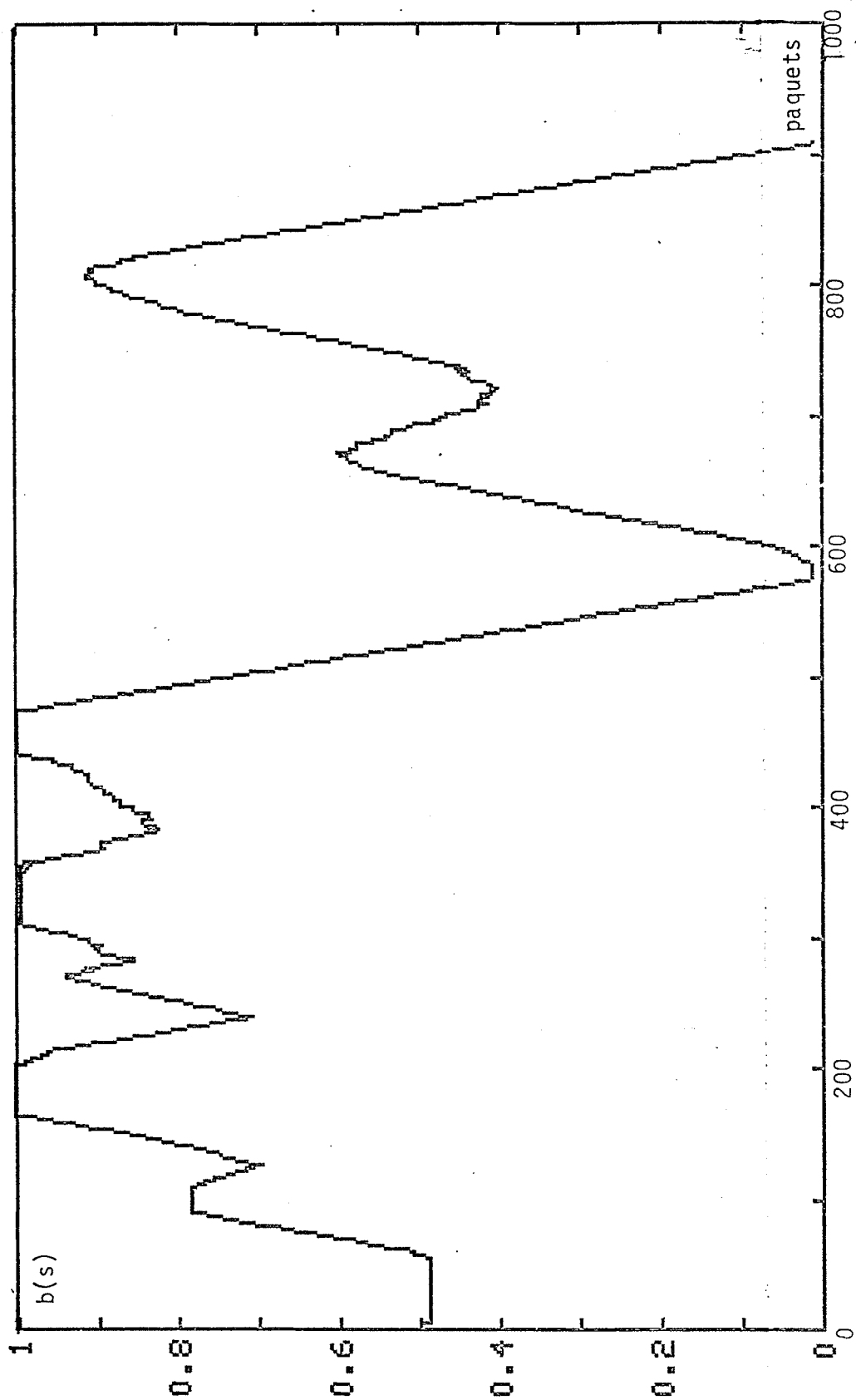


FIG 5-7 Variations du taux d'occupation du tampon d'entrée pour chaque paquet

Quant aux variations de la fondamentale, elles n'ont pas leur place comme tel puisqu'on prend la valeur de la variance tous les $60/b_i$ échantillons et cette valeur est alors utilisée pour le calcul de l'allocation. Cela donne à peu près le même effet que de calculer une variance instantanée et, aux fins de limiter l'information latérale, n'en prendre qu'une valeur toutes les m estimations, cette dernière valeur n'étant d'ailleurs pas nécessairement représentative des échantillons à venir.

Nous avons travaillé plusieurs semaines à essayer d'améliorer l'algorithme. Nous sommes même allé jusqu'à essayer des paquets de 6 bits (i.e. les mises à jour sont tellement nombreuses que l'information latérale représente alors le tiers de toute l'information transmise), et même dans ce cas, le rapport S/B segmentaire est de 20.4 dB (taux fixe= 20.6 dB). A partir du moment où l'on réalise que le seul gain possible est redevable aux périodes de la fondamentale, la réalisation pratique d'un quelconque algorithme de codage temporel à taux variable apparaît comme impossible.

VI.- CONCLUSION

Dans le présent chapitre, nous reprenons les conclusions auxquelles nous sommes arrivés au cours du projet, et ce, dans le but de faire plus de lumière sur le comportement du codeur à taux variable et les raisons pour lesquelles on obtient un rendement aussi faible.

6.1 DOMAINE DU TEMPS ET DOMAINE DE FREQUENCES

Dans le codage à taux variable que nous avons proposé au cours de ce projet, il s'agissait de coder un bloc de n échantillons répartis sur l'axe du temps, en attribuant à chacun un nombre de bits selon un critère prédéterminé. Nous concluons à ce sujet, qu'il n'est pas possible d'obtenir un gain net autrement qu'en profitant des variations rapides (dues aux périodes de la fondamentale) qui surviennent sur l'axe du temps. Pour profiter de ces variations rapides et en tirer un gain, il faut pour cela transmettre une information latérale à un taux beaucoup trop élevé pour être praticable.

On peut alors faire un rapprochement et demander pourquoi l'allocation à taux variable donne des résultats intéressants dans un contexte de codage par transformée [19,20]? La réponse en fait peut être donnée très simplement si on réalise que les variations rapides

dans le domaine du temps correspondent aux variations lentes dans le domaine des fréquences (ou par extension, dans le domaine de la transformée). En d'autres termes, la solution à notre problème, qui est de transmettre les variations temporellement rapides avec peu d'information latérale, réside dans une transformation du bloc de N échantillons en un bloc de M valeurs appartenant au domaine de la transformée et à partir de là, on peut opérer un lissage de l'énergie du spectre et ne transmettre que les points stratégiques de l'enveloppe. Après réception et décodage, on pourra alors exécuter une transformation inverse et récupérer le signal avec un rapport S/B accru par rapport au codage temporel à taux fixe. C'est finalement la façon dont on opère dans les schémas de codage par transformée [19,20].

6.2 CONCLUSIONS GLOBALES DU PROJET

Si on fait maintenant le bilan des conclusions auxquelles nous sommes arrivés au cours de ce projet, nous pouvons dégager celles-ci en plusieurs énoncés comme suit:

- 1) Le gain de tout algorithme de codage à taux variable provient des variations d'énergie à l'intérieur du bloc considéré. Ces variations augmentent le rapport entre la moyenne arithmétique et la moyenne géométrique des composantes du bloc (éq. 2.18 et 4.04).

Si aucune variation n'est présente, alors aucun gain n'est possible.

2) Il existe deux types de variations temporelles dans les séquences de parole: les variations "rapides" dues aux stries de la fréquence fondamentale (phonèmes voisés), et les variations "lentes" dues aux passage d'un phonème à l'autre, ou encore d'un phonème à un silence etc.. Ces deux types de variations amènent les deux montées rapides que l'on retrouve dans la courbe 2 du gain (fig 4-1) (Première montée entre 2 et 8 msec., seconde montée entre 64 et 256 msec.).

3) Alors que le codage à taux fixe vise un rapport S/B constant quelque soit l'énergie du signal d'entrée, le codage à taux variable vise, quant à lui, une distorsion fixe (éq. 2.15, 2.16) au prix d'un rapport S/B dont la valeur peut varier considérablement à l'intérieur d'un bloc (voir tableau VII section 4.2.3).

4) Les deux méthodes d'allocation: formule analytique (éq. 2.14) et gain marginal (éq. 3.10) sont équivalentes en pratique. La première méthode est rapide mais ne garantit pas le respect de la contrainte sur la somme des bits (éq. 2.01). La seconde méthode (gain marginal) est exacte et mieux appropriée à l'allocation de nombres entiers de bits et finalement plus adaptable à différentes répartitions de la probabilité des amplitudes.

- 5) En pratique, bien que la distorsion devienne plus "constante" lorsqu'on utilise une allocation variable, on est encore loin d'atteindre la limite théorique et ce, à cause de contingences telles que: l'allocation d'un nombre entier de bits et la nécessité que l'allocation soit positive ou nulle tout en étant limitée à une plage connue.
- 6) Du fait que l'oreille humaine a tendance à préférer un rapport S/B constant (ou qualité uniforme), le rapport S/B segmentaire représente une meilleure mesure du jugement subjectif ([32],[33], voir aussi section 4.3) puisqu'il est plus sensible aux variations de qualité. A cause de cela, il serait souhaitable d'optimiser le rapport S/B segmentaire. Malheureusement, nous ne connaissons pas de moyen simple de le faire. En effet, de par la nature de la formule de mesure segmentaire, on ne peut plus considérer les allocations d'une manière indépendante d'un échantillon à l'autre. On a pu toutefois prouver que le gain du rapport S/B standard représente une limite supérieure de ce qu'on pourrait éventuellement atteindre avec le gain du rapport segmentaire (section 4.2.3 et appendice B).
- 7) Le seul gain que l'on peut obtenir du point de vue segmentaire, est attribuable aux stries de la fondamentale. Du fait que ces stries représentent des variations rapides, l'oreille humaine "intègre" alors les variations du rapport S/B qui découle des allocations variables, et ne ressent que l'amélioration moyenne du S/B qui en

résulte.

Dans le cas des variations lentes (phonèmes, silences...), l'oreille humaine perçoit alors ces variations de rapport S/B d'un segment à l'autre et les ressent comme une diminution de la qualité globale par rapport à l'allocation à taux fixe. Ce comportement est fidèlement traduit par le rapport S/B segmentaire (courbe 3 de la figure 4-1).

- 8) Du fait que le seul gain net possible provient de la période de la fondamentale, une réalisation "pratique" d'un tel algorithme est rendue inopportune puisqu'alors le taux d'information latérale serait beaucoup trop important.
- 9) En voulant n'exploiter que les variations lentes du signal de parole (par l'utilisation d'une variance syllabique ou encore le lissage d'une variance instantanée) on obtient un résultat pire que lorsque l'allocation est fixe (Sections 4.2.2, 4.2.3 et 5.2.5).
- 10) La seule solution possible en pratique du moins, est d'opérer au préalable une transformée sur le bloc d'échantillons d'entrée. Une telle transformation amène alors une correspondance entre les variations rapides dans le domaine du temps avec les variations lentes

dans le domaine de la transformée, rendant ainsi possible un gain net. Cette méthode a d'ailleurs fait ses preuves dans les schémas de codage par transformée.

En somme, bien qu'en théorie le codage à taux variable paraisse très prometteur, il en va tout autrement en pratique. Ce n'est pas tant la théorie qui est en défaut, mais la mesure de qualité objective qui est utilisée (S/B standard) et sur laquelle se fonde toute la théorie. Si on ajoute à cela les contraintes propres à toutes réalisations "pratiques" (ex.: l'allocation doit être un nombre entier), le gain du codage à taux variable devient peu "rentable" en comparaison de la complexité requise et le délai.

Nous n'entrevoions donc pas, à court terme du moins, d'utilisation pratique du codage à taux variable tel qu'appliqué au domaine du temps.

BIBLIOGRAPHIE

Théorie de l'information ...

1. R.G. Gallager, "Information Theory and Reliable Communication", New York, Wiley 1968.
2. J.J.Y. Huang, P.M. Schultheiss, "Block Quantization of Correlated Random Variables", IEEE Trans. on Comm. Systems CS-11, pp. 289-296, Sept 1963.
3. A. Segall, "Bit Allocation and Encoding for Vector Sources", IEEE Trans. on Info. Theory IT-22 No 2, pp. 162-169, Mars 1966.
4. J. Max, "Quantizing for Minimum Distortion", IRE Trans. on Info. Theory, pp. 7-12, Mars 1960.
5. H. Abut, N. Erdöl, "Bounds of $R(d)$ Functions for Speech Probability Models.", IEEE Trans on Info. Theory IT-25 no 2, pp. 225-228, Mars 1979.

Traitement de signaux et codage numérique de la parole...

6. A.V. Oppenheim et R.W. Schafer, "Digital Signal Processing", Prentice-Hall 1975.
7. L.R. Rabiner et R.W. Schafer, "Digital Processing of Speech Signals", Prentice-Hall 1978.
8. P. Cummiskey, N.S. Jayant, "Adaptive Quantization in Differential PCM Coding of Speech", BSTJ vol 52, pp. 1105-1118, Sept. 1973.
9. P. Noll, "A Comparative Study of Various Quantization Schemes for Speech Coding", BSTJ vol 54, pp. 1597-1614, Nov. 1975.
10. N.S. Jayant, "Adaptive Quantization with a One Word Memory", BSTJ vol 52, pp. 1119-1144, Sept 1973.
11. N.S. Jayant, "Digital Coding of Speech Waveforms: PCM, DPCM and DM Quantizers", Proc. IEEE 62, pp. 611-632, Mai 1974.
12. R.E. Crochière, "A Midrise/Midtreast Quantizer Switch for Improved Idle Channel Performance in Adaptive Coders ", BSTJ vol 57, pp. 2953-2954, Oct. 1978.

13. P. Castellino, G. Modena, L. Nebbia, C. Scagliola, "Bit Rate Reduction by Automatic Adaptation of Quantizer Stepsize in DPCM Systems", 1974 Zürich Seminar, article B6.
14. J.B. Anderson, "Tree Coding of Speech", IEEE Trans. on Info. Theory IT-21, pp. 379-387, Juil. 1975.
15. J.B. O'Neal Jr., "Differential Pulse-Code Modulation (PCM) with Entropy Coding", IEEE Trans. on Info. Theory IT-22 No 2, pp. 169-172, Mars 1976.
16. J. Menez, F. Boeri, D.J. Estaban, "Optimum Quantizer Algorithm for Real Time Block Quantizer", IEEE 1979 Int. Conf. on ASSP, pp. 980-984.
17. J.A. Bucklew, N.C. Gallagher, "Some Properties of Uniform Stepsize Quantizers", IEEE Trans. on Info. Theory IT-26, pp. 610-613, Sept. 1980.
18. M.D.Paez, T.H. Glisson, "Minimum Mean Square Error Quantization in Speech PCM DPCM Systems.", IEEE Trans. on Comm. COM-20, pp. 224-230, Avril 1972.
19. R. Zelinski, P. Noll, "Adaptive Transform Coding of Speech Signals", IEEE Trans. on Acoustic, Speech and Signal Proc. ASSP-25, pp. 229-309, Aout 1977.
20. D. Sloan, "Adaptive Transform Coding of Speech", Rapport technique INRS Télécom., no 79-05, 1979.
21. B.S. Atal, M. R. Schroeder, "Predictive Coding of Speech Signals and Subjective Error Criteria", IEEE Trans. on Acoustics, Speech and Signal Proc. ASSP-27, pp. 247-254, Juin 1979.
22. J. L. Flanagan, M.R. Schroeder, B. S. Atal, R.E. Crochière, J. M. Tribolet, "Speech Coding", IEEE Trans. on Comm. COM-27, pp. 710-736, Avril 1979.

Codage à taux Variable...

23. J.J. Dubnowski, R.E. Crochière, "Variable Rate Coding of Speech", BSTJ vol. 58, pp. 577-600, Mars 1979, ainsi que: IEEE 1979 Int. Conf. on ASSP pp. 445-448.
24. R.V. Cox, R.E. Crochière, "Multiple User Variable Rate Coding for TASI and Packet Transmission Systems", IEEE Trans. on Comm. COM-28, pp. 334-344, Mars 1980.
25. A.G. Tescher, R.V. Cox, "Image Coding: Variable Rate DPCM through

Fixed Rate Channel ", Proc. Society of Photo-optical Instr. Eng.
no 119, pp. 147-154, Aout 1977.

26. C.J. Weinstein, E.M. Hofstetter, " The Trade-off Between Delay and TASI Advantages in a Packetized Speech Multiplexer", IEEE Trans. on Comm. COM-27, pp. 1716-1720, Nov. 1979.
27. D. Cohn, J. Melsa, "The Relationship Between Adaptive Quantizer and a Variance Estimator", IEEE Trans. on Info. Theory IT-21, pp. 669-670, Nov. 1975.

Mesures de qualité du signal de parole...

28. IEEE Recommended Practice for Speech Quality Measurements
IEEE Trans in Audio Electroacoustic, 17, pp. 225-246, Sept 1969.
29. D.C. Stevenson, "Procedure Programs for Preference Testing
of Speech Samples", Rapport interne BNR, Fév. 1981.
30. R.E. Crochière, L.R. Rabiner, N.S. Jayant, J.M. Tribolet, "A Study
of Objective Measures for Speech Waveform Coders", 1978 Int. Zürich
Seminar on Digital Comm., Mars 1978.
31. B. McDermott, C. Scagliola, D. Goodman, "Perceptual and Objective
Evaluation of Speech Processed by Adaptive Differential PCM",
BSTJ vol 57, pp. 1597-1618, Mai-Juin 1978.
32. P. Mermelstein, "Evaluation of Segmental SNR Measure as an
Indicator of the Quality of ADPCM Coded Speech", J. of Acoust.
Soc. of Am., no. 66, pp. 1664-1667, Déc. 1979.
33. M. Nakatsui, "Subjective Speech-To-Noise Ratio as a Single Measure
of Speech Quality for Digital Coders", 1980, (A être publié dans
J. of Acoust. Soc. of Am.).

Autres sujets...

34. D.W. Mitchler, M.I. Elmasry, "A C-MOS Digital Echo Cancellor
IEEE Conf. canadienne sur les télécommunications et l'énergie,
Montréal 1980, pp. 342-345.
35. D. E. Knuth, "The Art of Computer Programming; Sorting and
Searching", Addison-Wesley, 1973.
36. G. K. Helder, "Customer Evaluation of Telephone Circuits
with Delay", BSTJ vol 45, pp. 1157-1191, Sept. 1966.

APPENDICE A: GAIN THEORIQUE DU CODAGE
A TAUX VARIABLE.

Le but de l'appendice est de démontrer que le gain obtenu sur le rapport S/B dans le cas du codage à taux variable, est insensible aux variations des paramètres r et K dans l'équation générale suivante:

$$A.01 \quad d_i^2 = S_i^2 \cdot 2^{-r(b_i - K)}$$

sujet à la contrainte de l'équation 2.01:

$$A.02 \quad \sum_{i=1}^N b_i = N \cdot C$$

Avec le codage à taux fixe, il est aisé de montrer que le rapport S/B devient:

$$A.03 \quad S/B = 10r(C - K) \log_2$$

Dans le cas du codage à taux variable, par la méthode des multiplicateurs de Lagrange, tel que montré au chapitre II, on obtient:

$$A.04 \quad b_j = K - \frac{1}{r} \log_2 \left[\frac{N t}{r S_j^2} \right]$$

L'équivalent de l'équation 2.14 devient alors:

$$A.05 \quad b_j = C + \frac{1}{r} \log_2 \left[\frac{S_j^2}{(\prod_i S_i^2)^{\frac{1}{N}}} \right]$$

Le rapport S/B prend finalement la forme suivante:

$$A.06 \quad S/B = 10r(C-K) \log_2 + 10 \log \left[\frac{\frac{1}{N} \sum_i S_i^2}{(\prod_i S_i^2)^{\frac{1}{N}}} \right]$$

On observe donc que le gain, c'est-à-dire le second terme de l'équation A.06, est indépendant de r et K.

Dans le cas d'une distribution laplacienne, l'équation de d_i^2 en fonction de S_i^2 et de l'allocation b_i se conforme au modèle général de l'équation A.01 dans lequel $r=2$ (comme le cas gaussien), et $2^{-2K} = e/\pi$ [5]:

$$A.07 \quad d_i^2 = \frac{e}{\pi} S_i^2 2^{-2b_i}$$

La valeur du gain, exprimée par l'équation 2.18 ne change donc pas pour une distribution laplacienne, laquelle est beaucoup plus voisine du comportement du signal de parole.

APPENDICE B: COMPLEMENT SUR LE RAPPORT S/B SEGMENTAIRE

Dans l'appendice qui suit, nous examinerons la forme générale du rapport S/B segmentaire dans le cas du codage comportant une allocation fixe ou variable..

Posons: C.....taux du canal (bits/éch.)

M.....nombre de segments (128 éch. par segment)

N.....nombre d'échantillons dans un bloc

S_i^2valeur de l'estimation de variance du ième échantillon.

Soit aussi $\bar{S}_n^2$, la variance moyenne pour le nième segment, qu'on exprime comme:

$$\text{B.01} \quad \bar{S}_n^2 = \frac{M}{N} \sum_{j=(n-1)\frac{N}{M}}^{n\frac{N}{M}} S_j^2$$

On peut alors exprimer le rapport S/B segmentaire comme étant:

$$B.02 \quad S/B \text{ seg} = 10 \log \left[\prod_{n=1}^M \frac{\bar{S}_n^2}{\bar{d}_n^2} \right]^{1/M}$$

Nous allons maintenant examiner chaque cas de codage séparément:

a) Codage à taux fixe:

Dans ce cas, la distorsion (cas gaussien) s'exprime comme:

$$B.03 \quad d_i^2 = S_i^2 2^{-2(C-K)}$$

d'où il découle:

$$B.04 \quad \bar{d}_n^2 = \bar{S}_n^2 2^{-2(C-K)}$$

On obtient finalement, pour le cas du codage à taux fixe:

$$B.05 \quad S/B \text{ seg} \Big|_{\text{T.F.}} = 10 \log \left[\prod_{n=1}^M 2^{2(C-K)} \right]^{1/M} = 20(C-K) \log 2$$

On constate alors que le résultat obtenu est exactement le même qu'avec le S/B standard (éq. 2.06), ce qui nous amène à la constatation suivante:

Constatation #1:

Dans le cas du codage à taux fixe (ici nous parlons d'un cas "idéal"), les rapports S/B standards et segmentaires sont théoriquement égaux, ce qui est prévisible puisque le codage à taux fixe "idéal" réalise un rapport S/B constant d'un échantillon à l'autre.

b) Codage à taux variable

Dans ce cas, la distorsion est fixe (c.f. éq 2.16) et on obtient:

$$\text{B.06} \quad \bar{d}_n^2 = d_{\min}^2 = d^{*2} 2^{-2(C-K)} \quad \text{où} \quad d^{*2} = \left(\prod_{i=1}^N S_i^2 \right)^{1/N}$$

d'où on tire:

$$\text{B.07} \quad \left. \begin{array}{l} \text{S/B seg} \\ \text{I.V.} \end{array} \right| = 10 \log \left[\frac{\left(\prod_{n=1}^M \bar{S}_n^2 \right)^{1/M}}{d^{*2} 2^{-2(C-K)}} \right] = 20(C-K) \log 2 + 10 \log \frac{\left(\prod_{n=1}^M \bar{S}_n^2 \right)^{1/M}}{\left(\prod_{i=1}^N S_i^2 \right)^{1/N}}$$

On voit que l'équation B.07 ressemble beaucoup à l'équation 2.17 du S/B standard sauf pour le numérateur du terme d'extrême droite. Si on compare ces deux dernières équations, et qu'on introduit la relation suivante:

$$B.08 \quad \left[\prod_{m=1}^M S/B_m^{(128)} \right]^{1/M} \leq \frac{1}{M} \sum_{m=1}^M S/B_m^{(128)} = \frac{1}{N} \sum_{i=1}^N S/B_i$$

alors, on peut faire un second constat:

Constatation #2:

Le gain du rapport S/B segmentaire, de par sa nature, est toujours égal ou moindre que le gain du rapport S/B standard.

D'ailleurs, à ce sujet, on constate que si la mesure ne contient qu'un seul segment ($M=1$), on revient exactement à l'équation 2.17 pour le S/B standard. Dans le prochain énoncé, nous voulons souligner un fait important:

Constatation #3:

Lorsque le segment ne contient qu'un seul échantillon ($M=N$), le gain du codage à taux variable devient nul.

Pour s'en convaincre, il suffit de poser $M=N$ dans l'équation B.07 . Une autre manière de prouver ceci est de considérer le rapport S/B d'un échantillon comme étant une variable aléatoire S/B dont la moyenne est :

$$B.09 \quad E(S/B_i | b_i = J) = 6(J-K)$$

l'équation B.09 correspondant à B.05 pour lequel on fait l'approximation que $20\log 2$ vaut 6. De plus, on utilise la contrainte d'allocation :

$$B.10 \quad \sum_{i=1}^N b_i = N \cdot C$$

Ce qu'on cherche, c'est la moyenne de l'espérance du rapport S/B (en décibel) de chaque échantillon (ce qui équivaut à une mesure segmentaire où chaque segment dure un seul échantillon) :

$$B.11 \quad \frac{1}{N} \sum_{i=1}^N E(S/B_i | b_i)$$

Par les équations B.09 et B.10, on peut transformer B.11 et écrire :

$$B.12 \quad \frac{1}{N} \sum_{i=1}^N E(S/B_i | b_i) = \frac{1}{N} \sum_{i=1}^N G(b_i - K) = G(C - K)$$

le dernier terme étant obtenu de l'équation B.09.

Ceci prouve donc la constatation #3, à savoir que la moyenne des rapports S/B de chaque échantillon exprimés en décibels (éq. B.09) est égale à la valeur du rapport S/B standard (éq. 2.06).

APPENDICE C: COMPARAISON DE LA COMPLEXITE DES METHODES D'ALLOCATION

Notre but, dans cette appendice, est de détailler l'approche utilisée pour déterminer le nombre d'opérations requises pour effectuer l'allocation d'un bloc de N échantillons à un taux moyen de C bits.

a) Méthode analytique

Si on se reporte à la formule analytique pour l'allocation :

$$C.01 \quad b_i = C + \frac{1}{2} \log_2 \left(\frac{S_i^2}{d^{*2}} \right)$$

où

$$C.02 \quad d^{*2} = \exp \left[\left(\sum_{i=1}^N \log S_i^2 \right) \times N^{-1} \right]$$

Si on compte une opération unique pour l'addition et le logarithme, et quatre opérations pour les multiplications (Dépendant

de la méthode utilisée pour effectuer les multiplications, on peut compter différentes quantités d'opérations. Toutefois, nous nous en tenons ici à la méthode de comptage utilisée par A.V. Oppenheim et R.W. Schafer [6]), on obtient de l'ordre de $2N$ opérations pour C.02 et 10 opérations pour C.01. Comme l'équation C.01 doit être exécutée $N.I$ fois (I étant le nombre d'itérations requises pour obtenir une précision d'allocation), on obtient finalement:

$$C.03 \quad 10 \cdot N \cdot I + 2 \cdot N$$

b) Méthode de gain marginal

Il s'agit cette fois dans cette méthode d'évaluer le gain marginal P_{ij} pour les $C.N$ bits à allouer. Comme P_{ij} représente une seule multiplication (4 opérations), on arrive au total de $4.N.C$ opérations pour le calcul des P_{ij} .

La première série de N valeurs de P doit ensuite être classée par ordre et pour cela, suivant une méthode efficace (ex.: arbre binaire), il faut environ $3.N.\log N$ opérations. Une fois la première série de P_{ij} calculée et classée, on doit reclasser un nouveau P_{ij} à chaque allocation suivant cette fois une méthode par insertion. Comme la valeur de P_{ij-1} était la plus grande dans la liste, on peut supposer

que la nouvelle valeur de P_{ij} devrait se classer dans la première moitié de la liste, nécessitant ainsi en moyenne $N/4$ tests pour les $N.C$ valeurs, soit un total de $N^2.C/4$ opérations.

Le nombre total des opérations est donc pour la méthode de gain marginal:

$$C.06 \quad 4NC + 3N \log_2 N + \frac{N^2 C}{4} = O[N^2]$$